ESSAYS IN ENVIRONMENTAL ECONOMICS

by

Takuro Miyamoto

B.A., Waseda University, 2003
M.A., Keio University, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

September, 2011

Abstract

The objective of this dissertation is to improve our understanding of environmental policies, particularly with respect to two emerging alternative approaches to regulation. They are alternatives to the command and control approach, which policymakers have relied on heavily since the early 1970s. One alternative is to introduce market-based policy instruments like emission taxes or tradable permits. Another alternative is to rely on voluntary approaches to environmental protection. This thesis will study these two alternatives. Our first essay will focus on voluntary programs (VPs) that aim to reduce emissions of pollutants. We try to explain theoretically why governments implement these programs and to examine the property of the VP which the regulator implements to maximize social welfare. We show that if setting an efficient mandatory standard is politically difficult, a regulator might implement the VP because it can generate higher social welfare than the mandatory standard. The abatement rate of the VP that generates the highest social welfare costs participating firms the same amount as the mandatory standard would. The second essay will empirically examine the determinants of environmental management system certifications, especially the ISO 14001 certification, which is a popular environmental practice, and their impacts on environmental performance. In particular, we focus on intra-industry spillovers of ISO 14001 adoption and environmental performance. We apply estimation methods of spatial econometrics to a Japanese facility's dataset to deal with the spillovers. We find intra-industry spillovers of emissions reduction into the air between similarly sized facilities and of ISO 14001 adoption between similarly sized facilities that emit into water. The third essay will compare taxes and quotas, when an informed polluting industry influences them by political contributions to a government. We show that private information can improve social welfare under taxes but cannot improve it under quotas. Private information also reduces a comparative disadvantage of the taxes over the quotas when the government does not care about social welfare very much.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my gratitude to my supervisory committee, Brian Copeland, Sumeet Gulati and Werner Antweiler for their support at each stage of this thesis. I am grateful to Brian Copeland and Sumeet Gulati for their suggestions and encouragements, which led me to think more deeply about the subject and helped me complete the theoretical chapters of this thesis. In addition, I warmly thank Werner Antweiler for his valuable advice. His extensive discussions concerning my work have been very helpful for an empirical chapter of this thesis. I would also like to thank my external examiner Per Fredriksson for carefully reading this dissertation and making useful comments.

The support of fellow students and friends over the years is deeply acknowledged with appreciation. Finally, I also would like to thank my wife Noriko and my parents Yasuhide and Tsurue for their support over the years.

Chapter 1

Objective and Overview

1.1 Objective

In the 1960s and 1970s, the beginnings of the environmental movements in developed countries were associated with the establishment of environmental administrations and with the introduction and amendment of national environmental legislation. The approaches to pollution control during this period were mainly command-and-control and emission standard regulations. From the viewpoint of polluting firms, their environmental responsibility during this period meant regulatory compliance.

However, economists have urged the use of market-based instruments for pollution control instead of command-and-control regulations because command-and-control approaches are less efficient than market-based instruments and do not provide dynamic incentives for technological innovation and its diffusion. In addition, politicians and environmentalists recognized the need for cost-effective instruments because pollution control costs were increasing. Therefore, market-based instruments were introduced in the 1980s and 1990s. Tradable permit systems were introduced in the US, while environmental taxes were introduce in European countries.

More recently, voluntary approaches, commitments to improve environmental performance beyond legal requirements, have played a prominent role in addressing environmental challenges. One reason for the popularity of voluntary approaches from the supply side of environmental regulations is that it is politically difficult to pass environmental laws, especially in the US. However, affected firms started to tackle environmental issues not only because of regulation but also in light of potential cost reduction, liability concerns, and the indirect impact of regulations in the 1980s. Therefore, governments have offered voluntary environmental programs (VEPs) and encouraged firms to take voluntary environmental actions. For example, the US Environmental Protection Agency (EPA) has initiated over one hundred VEPs since the 33/50 Program was launched. Another example involves Japanese local governments that have provided support programs for the adoption of environmental management system (EMS) certificates, such as ISO 14001. These support programs may have contributed to the fact that the number of certifications in Japan was the highest in the world until 2008.

This thesis will examine environmental policies, voluntary approaches in particular in chapters 2 and 3, and emission quotas and market-based instruments in chapter 4. Voluntary programs for emissions reductions will be studied in chapter 2. Voluntary actions by Japanese facilities will be examined in chapter 3. In chapter 4, we compare emission taxes and quotas.

Chapter 2 uses a theoretical model to explain why governments implement voluntary emission reduction programs (VPs). The chapter also examines the properties of the VP that the regulator sets to maximize social welfare. In the model, there are 3 types of players: multiple polluting firms, a regulator, and a legislator. The regulator has two options to generate emissions reductions from firms: a mandatory standard and a VP. The mandatory standard can force all firms to reduce their emissions. However, the approval of the legislator who is affected by the lobby group of the polluting firms is necessary to implement it. However, the regulator cannot force firms to participate in the VP, but legislator approval is not necessary to implement it. We show that the VP can generate less social cost and more aggregate abatement than the mandatory standard. Therefore, the regulator implements the VP when it can generate higher social welfare than the mandatory standard. We also find that changes in parameters affect aggregate abatement under the VP more than under the mandatory standard because such changes affect the abatement rate of individual firms and the participation rate.

Chapter 3 empirically investigates reasons why facilities voluntarily acquire an environmental management system (EMS) certification and the impacts of EMS certification on environmental performance in Japan. We focus on ISO 14001 adoption and its impacts on environmental performance by Japanese facilities during 2001–2003. According to a survey of the Organisation for Economic Co-operation and Development (OECD), many facilities see the fact that similar facilities adopt environmental practices as an important motivation for adopting them. Thus, intra-industry spillover of the adoption of environmental practices is likely to exist. Using estimation methods of spatial econometrics, we estimate ISO 14001 adoption and its impacts on environmental performance while controlling for such spillovers and examining the magnitude of spillovers. We show findings indicating that facilities emitting pollution into water are more likely to adopt ISO 14001 iffacilities that belong to firms with similar revenue in the same industry adopt it and that the percentage change in emissions into air is correlated between similar-sized facilities in the same industry.

Chapter 4 compares taxes and quotas when a polluting industry with private information on emission abatement costs politically influences the taxes and quotas through lobbying activity. We employ an informed principal model developed by Maskin and Tirole (1992). We examine how taxes and quotas affect a polluting industry's incentive to influence regulation and how private information affects social welfare and the political influence of the polluting industry. We show that private information can improve the social welfare under taxes but cannot improve it under quotas. Private information also reduces a comparative disadvantage of taxes over quotas when the government does not care about social welfare.

1.2 Overview

This thesis is organized as follows. Chapter 2 presents a theoretical analysis of voluntary emission reduction programs. Chapter 3 examines the impacts of ISO 14001 on environmental performance and the intra-industry spillovers of ISO 14001 adoption and environmental performance in Japan. Chapter 4 compares taxes and quotas when a polluting industry with private information on abatement costs engages in lobbying. Chapter 5 concludes this thesis. Proofs of the propositions in Chapter 2 are given in Appendix A. Appendix A also includes arguments on the equilibrium abatement rate and the number of participating firms when the equilibrium number of participating firms derived in proposition 2 is not an integer. In appendix B, we show how to calculate some tax rate and contribution used in chapter 4, and we prove propositions of chapter 4.

Chapter 2

Why Regulators Adopt Voluntary Programs: A Theoretical Analysis of Voluntary Pollutant Reduction Programs

2.1 Introduction

Voluntary approaches to environmental protection have become prominent since the late 1980s. In the European Union (EU), the number of new voluntary agreements increased from 6 in 1981 to more than 45 in 1995 (OECD (1999)). There are more than 300 negotiated agreements between governments and polluting industries or firms in Europe. In the United States (US), Brouhle et al. (2005) identified over 50 voluntary programs at the federal level alone since 1991, when the 33/50 program[1], the first voluntary program, was launched. In Japan, over 30,000 negotiated agreements between local governments and polluters are in effect[2].

Voluntary approaches can take various forms. However, based on government involvement, we can categorize them into three types. The first type is a unilateral action that is initiated and carried out by firms and industries. The second form is a bilateral agreement between the government and

---

[1] The US EPA launched this program to reduce aggregate emissions of 17 chemicals by 33% in 1992 and by 50% in 1995, relative to the 1988 baseline. We will exaplain more in section 2. Or see Khanna Khanna (2007) for a more detailed review of the 33/50 program.

[2] The reason why over 30,000 agreements are being implemented in Japan is that most agreements have been between a firm and municipality. When a firm constructs (or extends) its facility, it concludes an agreement with a municipality at which the facility is (will be) located. This is a typical setting in which agreements are made. Please see Welch and Hibiki (2003) for details. In contrast to Japan, many agreements occur at the federal government or industrial level in Europe and the US. This is why the number of agreements or voluntary policies is so different between Japan and other developed countries.

industries or firms. Both the government and polluters actively set the target and other parameters for the approach. The final type and the focus of this chapter is a voluntary program (VP) designed by a government. The government sets the objectives of the program, and the firms have a right to decide whether they will participate in the program. Thus, participation in these programs is "voluntary".

Because firms can choose not to join a VP, it is subject to the free-rider problem, which means that the VP is implemented but not undertaken by all firms. Therefore, participation rate is an important criterion when we evaluate VPs. Participation rates for some VPs are very low. Participation rates for the US 33/50 program and the Canadian ARET program, whose goal is to reduce the release and/or transfer of chemicals, are 17.0% and 13.4%, respectively. Relative to its participation rate, the total release and transfer of participating firms in 1988 (baseline year) was high (62.5%)[3]. However, the program did not cover about 40% of total releases and transfers.

Why do VPs continue to be implemented even when participation rates are low? Political difficulty in passing environmental laws is likely to contribute to the increased popularity of voluntary approaches, and therefore, it must be one of the reasons for the continued implementation of VPs despite low participation rates. As an example, a carbon tax was proposed at the EU level but failed due to industrial lobbying. Just after this tax failed to pass in the EU, many EU countries adopted voluntary approaches to climate change. Thus, governments adopt voluntary approaches instead of mandatory regulations in response to political pressure from polluting industries. Moreover, when a pollution problem is not a hot political issue, it is politically difficult to implement effective policies. Participation rate of VPs would likely be low in such cases.

In addition to these political difficulties, governments might sacrifice high participation rates for the sake of effectiveness. A government might have to set low abatement rates for each participating firm to maximize participation, and, as a result, aggregate abatement may be low. Thus, a high participation rate does not always reflect the effectiveness of the VP, and we cannot evaluate a VP based on the participation rate alone. However, the abatement rate of each participating firm is not an appropriate solitary criterion for evaluating a VP either. We have to take into account both rates

---

[3] Many of ARET substances were not required to be reported to the National Pollutant Release Inventory (NPRI), which legally mandates public reporting. Therefore, data on these substances are not available. See Antweiler and Harrison (2007) for NPRI-recorded emissions of ARET-listed substances and ARET-participating firms' share of them.

for each participating firm when evaluating a VP. In addition, examining the relationship between participation and abatement rates in effective VPs provides useful information for the design of new VPs.

This chapter aims to explain why governments implement VPs even though their participation rates can be low and to examine the relationships between participation rates, requirements (abatement rates) and the efficiency of VPs. To do so, we built a model with three types of players: a regulator, legislator, and multiple polluting firms. In the model, the regulator is assumed to be benevolent because it attempts to maximize social welfare. Two regulation options, a mandatory standard and a VP, are available to the regulator to reduce emissions. For the regulator to implement the mandatory regulation, the legislator, who is influenced by a lobby group of polluting firms through political contributions, must approve it. This legislative process of setting mandatory standards is intended to capture the political difficulty involved in setting the mandatory policy, as discussed above. In general, the mandatory standard is inefficient. Hence, the regulator might have an incentive to offer the VP to achieve more efficient abatement allocation, although the regulator cannot force the firms to participate. In addition to the regulator, the individual firms might have an incentive to participate in the VP to save lobbying costs.

Along with a sharp increase in the number of voluntary approaches, there has also been an increase in research on these approaches. Several papers have theoretically investigated voluntary agreements (VAs), and Lutz et al. (2000) and Maxwell et al. (2000) studied unilateral actions[4]. Lyon and Maxwell (2003) examined the effects of policy choice, taxes or VAs on the adoption of new technology when firms unilaterally adopt clean technology. Because governments design VPs but do not design the unilateral actions by polluters, the models and the results of these studies on unilateral actions cannot be applied to VPs. Most papers on VAs have built and analyzed single polluter models (e.g., Segerson and Miceli (1998), Hansen (1999), Segerson and Miceli (1999),

---

[4] Following the publication of the seminal papers of Segerson and Miceli (1998), Hansen (1999), and Segerson and Miceli (1999) who analyzed VAs, Manzini and Mariotti (2003) studied VAs between a regulator and group of heterogeneous firms. Glachant (2007) examined Vas, which cannot enforce a firm's (or an industry's) commitments by a model with a polluting industry, regulator, and legislator. Lutz et al. (2000) studied the effect of a "green" commitment by firms when their actions influence future regulation. Maxwell et al. (2000) showed the possibility of self-regulation preempting future regulation. The model of Lutz et al. (2000) and Maxwell et al. (2000) and its implications were explained Lyon and Maxwell (2004). See also Lyon and Maxwell (2002) for a survey of early theoretical and empirical studies of voluntary approaches.

and Glachant (2007)). Even with multiple polluters, all firms are assumed to collectively negotiate the VA with the government, as in Manzini and Mariotti (2003), or the voluntary agreement is not a subsidy for the adoption of technology (Lyon and Maxwell (2003)). Thus, most theoretical papers have not incorporated an individual firm's participation decision into voluntary programs or agreements with emission reduction targets, and, therefore, they have not explained why VPs with low participation rates are implemented.

The work of Dawson and Segerson (2008) is an exception in the literature. Dawson and Segerson (2008) developed models incorporating the participation of individual firms into the voluntary program or agreement. They examined the case in which the regulator has two options to achieve an (exogenous) aggregate emission level target via a VP or an emission tax. They analyzed the existence and properties of an equilibrium participation rate and its social welfare implications. However, there are no reasons that the regulator would offer a VP because it is always "not better" than the tax, and the regulator can always choose the tax. In contrast, the VP is implemented because the regulator preferred it in this chapter because of the presence of two public agents, a benevolent regulator and rent-seeking legislator, similar to the model described by Glachant (2007). If the benevolent regulator sets a mandatory policy, such as an emission tax, the policy is cost-effective. Therefore, the regulator has no incentive to implement the VP. However, the VP might generate higher social welfare than the mandatory policy if the rent-seeking legislator is the entity that would set the mandatory policy.

We show that the regulator can implement the VP, which generates less social cost and more aggregate abatement than the mandatory standard. By adopting the VP, the regulator can make participating firms allocate resources that would otherwise be used to lobby for emissions abatement. The regulator's problem is maximizing the resources that are reallocated from lobbying efforts to emissions abatement, subject to the constraint that there is no new participation if a portion of the abatement from the participating firms is greater than the aggregate abatement under the mandatory policy. In this case, the resource reallocated to abatement is maximized by making the "newest" participating firm abate as much as possible and putting the participation rate aside. Thus, in the most efficient VP, there are some "non-participating" firms. Moreover, we show that a VP with a low participation rate can generate less social costs and more aggregate abatement than the mandatory standard if it is politically difficult to set the mandatory standard.

The other results of this study are as follows. Changes in parameters affect aggregate abatement under the VP through two channels, the participating rate and the individual participating firms' abatement rates, but aggregate abatement under the mandatory policy is affected by the abatement rate of the individual firms alone. Therefore, changes in parameters affect aggregate abatement levels under the VP more than they do under the mandatory standard. If the mandatory policy is stringent, the VP is effective and its participation rate is high. For instance, when there are many polluters or each polluter has high emissions, the mandatory policy is stringent because aggregate emissions are large in such cases. Therefore, firms in a more pollution-intensive industry are more likely to join such a VP. The participation rate is also high when there are many polluters (if other parameters are the same). However, firms in an industry with higher abatement costs are less likely to join because the mandatory policy is less stringent.

Although the regulator implements a low participation rate VP when it is politically difficult to implement a mandatory policy, it is also possible that most or all firms will participate in the VP. Many firms participate in the VP when its net benefit is greater than that of the mandatory standard. These results are totally different from that of international environmental agreement (IEA) in Barrett (1994) who argued that polluters are free to join IEAs or not, much like a VP. The mandatory standard, a kind of punishment for non-participation to all polluters, creates the difference between VPs and IEAs.

This chapter is organized as follows. In next section, we review facts on participation in voluntary approaches. Section 3 presents a model and analyzes situations of equilibrium. We present a comparative statics analysis in Section 4. Section 5 discusses the results of changing the settings from a basic setting. Section 6 concludes this chapter.

2.2 Some facts on voluntary programs

In this section, we review some facts on participation in voluntary approaches based on the 33/50 program, the ARET program and environmental agreements in Europe. We then review the political

Table 2.1: Overview of 33/50 program
(Source: US Environmental Protection Agency (1999))

	1st round	2nd round	3rd round	Not contacted	Contacted	total
	Number of companies					
Companies with 33/50 facilities	504	4534	2512	2612	7550	10167
Participating companies (PCs)	328	819	140	7	1287	1294
Participation rate	65.1%	18.2%	5.6%	0.3%	17.2%	12.7%
	Quantities of 17 chemicals (millions of pounds)					
Total releases & transfers in 1988	993	367	45	89	1405	1496
Total from PCs	809	110	14	1	933	935
Share of releases & transfers from PCs	81.5%	30.0%	31.1%	1.1%	66.4%	62.5%

process of adopting voluntary approaches. The US 33/50 program was launched by the Environmental Protection Agency (EPA) in 1991. The goal of this program was to reduce the release and transfer of 17 chemicals by 33% in 1992 and by 50% in 1995, relative to the 1988 baseline. The EPA sent letters to the CEOs of the parent companies of the emitting facilities to encourage participation. In January 1988, the “top” 600 companies, which accounted for 66% of the total release of these chemicals in 1988, were invited to participate, and 5,400 other companies were invited in July 1991. After 1992, the EPA invited 2,512 companies that began emitting the aforementioned 17 chemicals after 1988 to participate. Table 2.1 gives an overview of the 33/50 program. The participation rate of the invited companies is 17.0%, and the share percentage of the total release and transfer of all substances by the firms invited in 1988 was 62.5%. Clearly, the participation rate for the first round of invited companies was much higher (65.1%) than it was for the companies invited later.

There have been several empirical studies on the 33/50 program. For example, Gamper-Rabindran (2006) found that plants in the chemical industry, which uses hazardous air pollutants intensively, were more likely to participate. She also found that the participation rate of firms in the chemical industry increased when the past inspection rate was high. Innes and Sam (2008) found that large firms listed as potentially-responsible parties for a large number of Superfund sites were more likely to join.

The Canadian Accelerated Reduction/Elimination of Toxics (ARET) program was launched in 1994. ARET targeted 117 toxic substances, including 30 substances that persist in the environment and may accumulate in living organisms. ARET required all firms to voluntarily reduce their release of those 30 substances by 90% and 87 other substances by 50% by the year 2000. This program was guided by a committee of stakeholders established in 1992, which included representatives from industry, health and professional associations, in addition to representatives from the federal and provincial governments. Although the ARET program has some characteristics of a negotiated voluntary agreement, it has more in common with the 33/50 program than with negotiated agreements because it called on all firms to reduce their emissions.

Antweiler and Harrison (2007) found that (1) large and emission-intensive facilities were more likely to participate; that (2) participation rates increased along with the participation of other facilities in the same industry (spillover within industry); and that (3) the participation rate for members of the trade associations that were involved in negotiating the terms of ARET was higher than for other facilities.

The participation of firms can be an important issue for negotiated voluntary agreements and for public voluntary programs. The European Environment Agency (1997) surveyed voluntary agreements in Europe and found that many small and medium enterprises (SMEs) tended to be free-riders when trade associations were directly involved with the agreements. In addition, the European Environment Agency (1997) admitted that free-riding by SMEs might jeopardize agreements even though it may not affect their efficiency. Hence, voluntary approaches to controlling pollution from multiple firms are subject to the problem of free-riders, which is one of the important issues to be addressed.

Political factors play a role in the adoption of voluntary approaches to environmental protection. Many voluntary programs and agreements have been adopted without approval from a legislative sector. For example, the EPA founded the 33/50 program, and the program was implemented without the approval of the legislative sector. In addition, in response to political pressure from polluting industries, some voluntary approaches have been adopted as an alternative to mandatory policies. For example, many EU countries adopted VAs in the mid-1990s after the EU carbon tax project was abandoned due to pressure from lobby groups for the energy-intensive industries.

Another example is the Clinton administration's Climate Change Action Plan (CCAP). Although the administration originally proposed an energy tax, it adopted voluntary programs to meet the goals of the CCAP due to political opposition from industries. Our model includes two types of public agencies, a legislative sector and a regulator, to capture this type of political process.

2.3 The model

The following three types of players are involved in our model: a regulator, a legislator and N polluting firms. The regulator is benevolent in that he/she aims to minimize social costs, the sum damage by pollution, and aggregate abatement costs. In the first stage, the regulator announces a voluntary $\alpha(\%)$ emission reduction program (VP). In the second stage, polluting firms decide whether to participate in the VP or not. Thus, the firms do not have to participate in the VP, but the regulator asks the legislator to enact a mandatory emission reduction standard if an insufficient number of firms participate. When the mandatory reduction standard is set or the insufficient number of firms participate in the VP, polluting firms will collectively lobby against it and then the legislator will set the mandatory standard. Therefore, the mandatory standard will typically be socially inefficient. On the other hand, if an sufficient number of firms participate in the VP in the second stage, then the VP is implemented. The timing of this model is described in Figure 2.1.

It is possible that the polluting firms will take voluntary action before the regulator announces the VP. However, if many firms are involved, it is hard for trade associations or firms to coordinate voluntary emission reduction efforts. This chapter considers such a case. Usually, polluting firms unite in opposition to mandatory regulations, although not all firms join such collective actions. In such a situation, the polluting firms form a lobby group, which will influence mandatory regulations through political contributions financed by voluntary contributions from the firms.

This chapter does not address the possibility that a legislator sets a mandatory policy even if a VP is already in place. However, to create a law that enforces the mandatory policy, various kinds of information, such as scientific knowledge on pollution, must be obtained and integrated. In many cases, it is likely logistically difficult and/or costly to obtain such technical information and expertise without the cooperation of a government-run environmental organization, which is represented

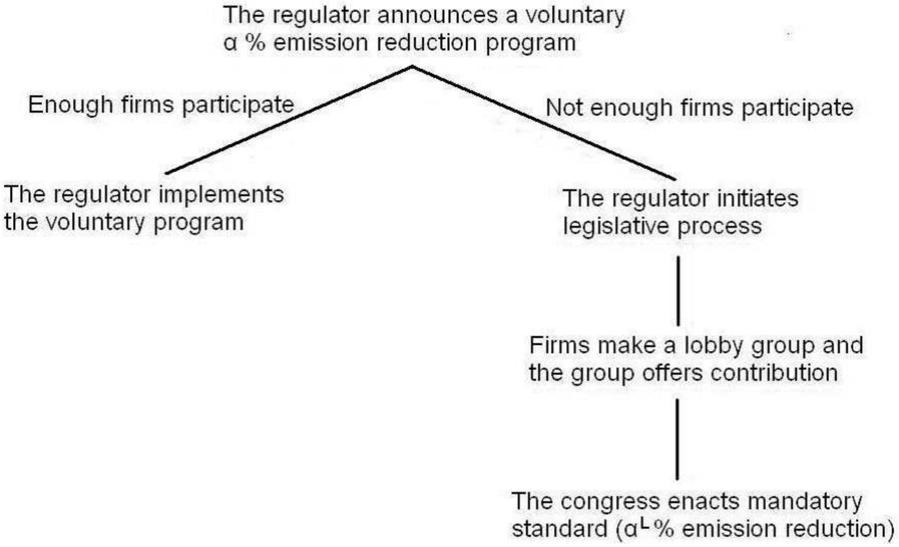


Figure 2.1: Decision tree of game

by the regulator in this chapter. Therefore, it is not unreasonable to assume that the regulator would initiate both policies.

Figure 2.1 shows that the regulator has two options: implement the mandatory policy or the voluntary program to minimize social costs, which are the sum of damage by pollution and aggregate abatement costs. We assume that the damage is a quadric function of emission, $\frac{1}{2}d(\sum_i(1-\alpha)\bar{e}_i)^2$, and that the abatement costs of a firm i is linear in emission, $c\alpha\bar{e}_i$ where $\bar{e}_i$, α, c and d are the pollutant emission of firm i before regulation, the reduction rate, marginal abatement cost and the slope of marginal damage, respectively. Because the regulator is benevolent, the regulator's cost under the mandatory policy is the sum of the aggregate abatement costs and damages

$$R^M(\alpha) = \frac{1}{2}d\left\{\sum_i^N((1-\alpha)\bar{e}_i)\right\}^2 + \sum_i^N c\alpha\bar{e}_i. \qquad \text{(Madantory policy)}$$

If the regulator could set a mandatory standard or a reduction rate, he/she would set the reduction rate

$$\alpha^* = 1 - \frac{c}{d\bar{e}_i N}.^5 \tag{2.1}$$

---

[5] We have this from the F.O.C for minimization of $R^M(\alpha)$. We focus on a case that (2.1) is nonnegative. Otherwise, we do not need to implement regulation.

However, the legislator sets the reduction rate for the mandatory program. When the congress enacts the mandatory standard, the lobby group of polluting firms influences it by offering political contributions to the legislator, who is a representative or median legislator. As Glachant (2007) described, the legislator's payoff function is assumed to be a weighted sum of political contributions, Ω and social welfare (negative social costs),

$$L(\alpha,\Omega)=\lambda\Omega-(1-\lambda)[\frac{1}{2}d(\sum_{i}^{N}(1-\alpha)\bar{e}_i)^2+\sum_{i}^{N}c\alpha\bar{e}_i]$$

where $\lambda\in(0,1)$. λ can be interpreted as the responsiveness of the legislator to lobbying or the political difficulty of setting an efficient mandatory abatement rate. Political difficulty may occur when a pollution problem is not an important issue on the political agenda. A high λ may reflect that it is not an important issue on the agenda.

Political contributions from the lobby group are financed by voluntary contributions from individual firms. Each firm is assumed to make a contribution to the lobby group taking other firms' contributions as given like voluntary public good provision games. Firm i's cost is the sum of abatement costs, $c\alpha\bar{e}_i$, and contributions to the lobby group, Ω^i, $c\alpha\bar{e}_i+\Omega^i$.

In contrast to the mandatory policy, the regulator can set the abatement rate of the VP, but it cannot force firms to participate. Hence, if the regulator implements the VP, the objective is

$$\min_{\alpha}R^V(\alpha)=\frac{1}{2}d\left\{\sum_{i}^{N_P}((1-\alpha)\bar{e}_i)+\sum_{j}^{N-N_P}\bar{e}_j\right\}^2+\sum_{i}^{N_P}c\alpha\bar{e}_i \qquad \text{(VP)}$$

where N_P is the number of firms participating in the VP. As mentioned above, individual firms can decide whether they participate in the VP. Firms have no incentive to join the VP if their cost under the VP is greater than it would be under the mandatory policy. Therefore, N_P and the VP participation rate depend on the reduction rate of the VP and on each firm's costs or reduction rate under the mandatory standard. In the next subsection, we analyze how the mandatory policy is set. Then, we examine how the regulator sets the VP.

2.3.1 The legislative subgame

First, we analyze the case in which the regulator uses the government to set the mandatory standard. In the final stage, the legislator sets the mandatory reduction rate, and firms reduce their emissions. Before the final stage, the lobby group offers political contributions depending on the reduction rate α. Because the legislator can reject the offer from the lobby group and choose a socially optimal reduction rate, the potential political contribution must satisfy $L(\alpha,\Omega) \geq L(\alpha^*,0)$. This constraint must hold equally because the total cost to the firms increases with increases in the reduction rate, and the lobby moves first to make a non-negotiable offer. Therefore, political contributions and the reduction rate have the following relationship

$$\Omega(\alpha) = \frac{1-\lambda}{\lambda}\left[\frac{1}{2}d(\sum_i(1-\alpha)\bar{e}_i)^2 + \sum_i c\alpha\bar{e}_i - [\frac{1}{2}d(\sum_i(1-\alpha^*)\bar{e}_i)^2 + \sum_i c\alpha^*\bar{e}_i]\right]. \tag{2.2}$$

Let $\hat{\alpha}(\Omega)$ be the reduction rate satisfying (2.2) when political contribution is Ω. Using $\hat{\alpha}(\Omega)$, we can describe an individual firm's problem. Their contributions affect political contributions directly and reduction rates indirectly. Therefore, taking it as a given that other firms have also made contributions, firm i chooses its contribution to minimize its total cost

$$\Omega^i + c\hat{\alpha}(\Omega^i + \sum_{j\neq i}\Omega^j)\bar{e}_i, \tag{2.3}$$

subject to (2.2). Using $\Omega = \Omega^i + \sum_{j\neq i}\Omega^j$, we enter (2.2) into (2.3) and take the derivative of (2.3) to yield

$$c\bar{e}_i\frac{\partial\hat{\alpha}}{\partial\Omega^i} = \frac{1-\lambda}{\lambda}[d\sum_i\bar{e}_i(\sum_i(1-\alpha)\bar{e}_i) - \sum_i c\bar{e}_i]\frac{\partial\hat{\alpha}}{\partial\Omega^i}. \tag{2.4}$$

Therefore, the abatement rate chosen via the legislative process is

$$\alpha^L = 1 - \frac{c}{dN\bar{e}} - \frac{c\lambda}{dN^2\bar{e}(1-\lambda)}.^6 \tag{2.5}$$

---

[6] In the following analysis, we assume α^L is positive. Abatement rate α^L should be 0 if α^L is negative. A contribu-

The third term of RHS is the difference in the reduction rates between the socially optimal and the mandatory standard. The absolute value of this difference decreases with the number of polluting firms because political contribution decreases due to free-riding. Therefore, if the socially optimal reduction rate is the same, the more firms pollute and the more effective the mandatory standard is.

If we focus on a symmetric equilibrium, then firm i's contribution is

$$\begin{aligned}\Omega^i &= \Omega(\alpha^L)/N \\ &= \frac{1-\lambda}{\lambda}\left[\frac{1}{2}d(\sum_i(1-\alpha^L)\bar{e}_i)^2+\sum_i c\alpha^L\bar{e}_i-[\frac{1}{2}d(\sum_i(1-\alpha^*)\bar{e}_i)^2+\sum_i c\alpha^*\bar{e}_i]\right]/N \\ &= \frac{c^2\lambda}{2dN^3(1-\lambda)} \end{aligned} \tag{2.6}$$

Finally, an individual firm's cost under the mandated standard is

$$\bar{C} = c\alpha^L\bar{e}+\Omega^i = c\bar{e}-\frac{c^2}{dN}-\frac{c^2\lambda}{dN^2(1-\lambda)}+\frac{c^2\lambda}{2dN^3(1-\lambda)}. \tag{2.7}$$

2.3.2 The VP subgame

If polluting firms participate in the VP, then their costs must be smaller than they would be under the mandatory standard. Hence,

$$c\alpha\bar{e}_i \leq \bar{C}. \tag{2.8}$$

Any participating firm will not change its participation decision in a state of equilibrium. Each firm makes the decision to participate or not, and therefore firms do not unilaterally change their participation in a state of equilibrium. Dawson and Segerson (2008) showed that this condition is equivalent to the condition that no participating firm has an incentive to unilaterally become a

---

tion is given by

$$\Omega_0 = \frac{1}{2}dN^2\bar{e}_i^2-Nc\bar{e}_i+\frac{c^2}{2d}.$$

If $\bar{C}_0 = \Omega_0$, then analysis of VP game is the same as in the case where α^L is non-negative.

non-participating firm and, technically

$$\frac{d}{2}[(N-(N_P-1)\alpha)\bar{e}_i]^2+(N_P-1)c\alpha\bar{e}_i \geq R^M(\alpha^L). \tag{2.9}$$

This condition means that the regulator will not implement the VP if only $N_P - 1$ polluting firms participate. If this condition does not hold, then the regulator will still implement the VP even if one of the participating firms does not participate (i.e., deviate). Thus, participating firms have an incentive to unilaterally deviate if this condition does not hold.

Subject to (2.8) and (2.9), the regulator tries to maximize its payoff or minimize social costs. Let α^V be such that $c\alpha^V\bar{e}_i = \bar{C}$. When the reduction rate under VP is α^V, the cost of the VP for the participating firms is the same as the costs of the mandatory policy. Then, the following lemma holds.

Lemma 2.1. *The regulator's problem is the same as maximizing aggregate abatement subject to* (2.8) *and* (2.9). *In addition, we can rewrite* (2.9) *as*

$$(N_P-1)\alpha\bar{e}_i \leq N\alpha^L\bar{e}_i. \tag{2.10}$$

Proof: Due to the linearity of abatement costs, the allocation of abatement does not affect social welfare. Based on (2.7) and (2.8),

$$\alpha^V = 1-\frac{c}{dN\bar{e}}-\frac{c\lambda}{dN^2\bar{e}(1-\lambda)}+\frac{c\lambda}{2dN^3\bar{e}(1-\lambda)} = 1-\frac{c}{dN\bar{e}}-\frac{c\lambda}{2dN^3\bar{e}(1-\lambda)}(2N-1) \leq 1-\frac{c}{dN\bar{e}_i} = \alpha^*.$$

Hence, under the VP, each participating firm abates less than the social optimal level, and the greater aggregate abatement results in a smaller social cost. Therefore, the regulator's problem is the maximization of the aggregate abatement, which is subject to (2.8) and (2.9), and (2.9) is equivalent to (2.10) . *QED*

The regulator's problem is the maximization of the aggregate abatement level, which is subject to (2.8) and (2.10). (2.10) implies that $(N_P - 1)$ participating firms' abatement levels must be below $N\alpha^L\bar{e}_i$ and independent of N_P. Hence, maximizing aggregate abatement is equivalent to maximizing the "N_P th" participating firm's abatement, and the following proposition holds.

Proposition 2.1. *The largest aggregate abatement under the VP is generated by the abatement rate* α^V *and the participation rate,* N_P^V/N*,*

$$N_P^V/N = \begin{cases} \alpha^L/\alpha^V + 1/N & \text{If } N > \alpha^V/(\alpha^V - \alpha^L) \\ 1 & \text{Otherwise.} \end{cases}$$

Proof: If $(N-1)\alpha^V\bar{e}_i \leq N\alpha^L\bar{e}_i$ or $N \leq \alpha^V/(\alpha^V - \alpha^L)$, then all firms have an incentive to join the VP because (2.8) and (2.10) hold for any $N_P \leq N$. Because α^V is the highest reduction rate, $\alpha^V$with $N_P^V/N = 1$ gives the highest aggregate abatement. If $N > \alpha^V/(\alpha^V - \alpha^L)$, (2.10) must hold with equality to maximize the aggregate abatement. So, By substituting $(N_P^V - 1)\alpha\bar{e} = N\alpha^L\bar{e}$ into $N_P^V\alpha\bar{e}$, we have $N_P^V\alpha\bar{e} = \alpha\bar{e} + (N_P^V - 1)\alpha\bar{e} = \alpha\bar{e} + N\alpha^L\bar{e}$. Because $N\alpha^L\bar{e}$ is set by the congress, the maximization of $N_P^V\alpha\bar{e}$ is equal to that of $\alpha\bar{e}$ or α. Therefore, the reduction rate which generates the largest aggregate abatement is the highest α that satisfies (2.8) and which is α^V. In addition, we have $N_P^V = N\alpha^L/\alpha^V + 1$ from (2.10). *QED*

See Appendix A.1 for the case where the N_P^V is not an integer. Two factors can be manipulated to increase the aggregate emission abatement or decrease social costs: the participation rate and the reduction rate. The regulator prefers high abatement and participation rates, but a trade-off exists between the reduction rate and the participation rate due to Constraint (2.10) which is a kind of a constraint on aggregate abatement. (2.10) means that the abatement by $N_P - 1$ firms must not be greater than the aggregate abatement level under the mandatory policy regardless of the participation rate. By substituting $(N_P - 1)\alpha\bar{e}_i = N\alpha^L\bar{e}_i$ ((2.10) with equality) into $N_P^V\alpha\bar{e}$, we have $N_P^V\alpha\bar{e} = \alpha\bar{e} + N\alpha^L\bar{e}$. Because $N\alpha^L\bar{e}$ is set by the congress, the largest aggregate abatement is achieved by choosing the highest reduction rate such that (2.8) holds.

This proposition means that the VP is the most environmentally effective if the regulator chooses the highest abatement rate, such that the participating firms' abatement costs under the VP are less than under the mandatory policy. If the regulator does choose the highest abatement rate, then some firms generally do not join the VP. Thus, symmetric firms may take asymmetric actions (i.e., some of them participate, but others do not). Of course, all the firms might participate in some cases[7]. In addition, the proposition implies that the regulator might implement the VP even if the participation rate is low (e.g., when the legislator does not care about the social costs related to pollution or when it is politically difficult to set the mandatory policy). In the next section, we discuss how changes in parameters (reduction rate, participation rate and aggregate abatement) affect the VP when it is effective.

To understand why symmetric firms might take asymmetric actions and what happens as a result, we consider simple cases with two firms. The payoffs for the firms and the regulator are described in Table 2.2. The left, center and right values in each cell represent the payoffs of Firm 1, Firm 2 and the regulator, respectively. P and NP stand for “participate” and “not participate”, respectively. The payoffs for firms in the case when both firms do not participate (NP, NP) are equal to the payoffs when the mandatory policy is implemented because the regulator prefers the mandatory policy in such a case. However, the firms’ payoffs in the other cases are the same as they are when the VP is implemented. The set of payoffs is likely to be similar to those in the left matrix when the legislator is not concerned with social welfare, which is low under the mandatory policy is low when the legislator cares less about it than about political contributions. However, payoffs are similar to those in the right matrix when the legislator cares about social welfare. It should be noted that a participating firm receives the same payoff under the VP as it would under the mandatory policy if the reduction rate is α^V(from (2.8)) .

Table 2.2: Payoff matrices when some firms participate (left) and when all firms participate (right).

		firm2	
		P	NP
firm1	P	0,0,1	0,1,0.5
	NP	1,0,0.5	0,0,0

		firm2	
		P	NP
firm1	P	0,0,0.5	0,1,-0.5
	NP	1,0,-0.5	0,0,0

Consider (P, NP) in the left matrix. Because the regulator’s payoff is greater than it would be in (NP, NP), the regulator will implement the VP. Given the participation of Firm 2, Firm 1 has no incentive to change from P to NP. In addition, Firm 2 has no incentive to change from NP to P.

---

[7] For example, the legislator cares about social welfare very much, or damage by pollution is serious.

Hence, (P, NP) is a VP equilibrium. (NP, P) is also a VP equilibrium because the same argument holds.

In the right matrix, neither (P, NP) nor (NP, P) is a VP equilibrium because the regulator prefers (NP, NP) over (NP, P) or (P, NP). Hence, the regulator will implement the mandatory policy, and the payoffs would be the same as with (NP, NP) if either firm does not participate. In addition, all players weakly prefer (P, P) over (NP, NP), and so they prefer (P, P). Therefore, neither firm has an incentive to change from P to NP. (P, P) is an equilibrium.

However, when no firms participate in the VP (NP, NP) and the mandatory policy is implemented, there is an equilibrium in both examples because the participation of one firm is irrelevant if the other firm does not participate. In the right matrix case, there also exists an equilibrium where either firm joins the VP, but the mandatory policy is implemented, as discussed above. As we can guess, "participation" versus "non-participation" is irrelevant to a firm if less than $N_P - 1$ of the other firms participate. Hence, there also exists a situation of mandatory regulation equilibrium.

In the two examples described above, the regulator's payoff under the VP equilibrium is the greatest of all the situations of equilibrium. Because the regulator is benevolent, this means that the VP can generate higher social welfare than the mandatory policy. Moreover, there is a situation of equilibrium where the mandatory policy is implemented. The next proposition formally shows these two outcomes.

Proposition 2.2. *The reduction rate* α^V *with participation rate* N_P^V/N *generates higher aggregate abatement and lower social cost than a mandatory standard. However, there also exist equilibria where only less than* $(N_P^V - 1)$ *firms participate in the VP and the mandatory policy is implemented.*

Proof: When $N_P^V < N$, $N_P^V \alpha^V \bar{e}_i > (N_P^V - 1)\alpha^V \bar{e}_i = N\alpha^L \bar{e}_i$. If $N_P^V = N$, $N_P^V \alpha^V \bar{e}_i > N\alpha^L \bar{e}_i$ because $\alpha^V > \alpha^L$. Lemma 1 implies that a higher aggregate abatement results in lower social costs. Therefore, α^V with N_P^V/N generates higher aggregate abatement and lower social cost than a mandatory standard.

However, the mandatory policy generates a higher aggregate abatement if the abatement rate is α^V and there is no participating firm, for example. The mandatory policy is implemented in such a case.

Consider Firm k's participation decision. We suppose that Suppose $N_P^V > 2$ and that other firms do not participate in the VP. It should be recalled that $(N_P^V - 1)\alpha^V \bar{e}_i = N\alpha^L \bar{e}_i$. The mandatory policy is implemented whether Firm k participates or not because the participating firms number fewer than $(N_P^V - 1)$ under both cases (firm k participate or it does not). Thus, firm k has no incentive to change its participation decision, from "not participate" to "participate". Therefore, there exists an equilibrium where no firms participate in the VP and the mandatory policy is implemented. Please see appendix A.2 for the proof of the general case. *QED*

When the mandatory policy is implemented, the firms' cost is the sum of the emission abatement costs and political contributions. Given that the firms cannot influence the VP through contributions, they do not spend resources lobbying for the VP. Therefore, the regulator can design the VP, which makes individual firms increase expenditures and reduce the amount of emissions relative to the mandatory policy (subject to (2.8) and (2.10)). It is possible that the VP can generate higher social welfare than the mandatory policy if enough firms join the VP.

In contrast with Dawson and Segerson (2008) model, the regulator in this model might prefer the VP over the mandatory standard if the participation rate is high enough, which depends on the firms' beliefs about other firms' participation decisions. This result occurs due to the difference between the regulator's objective and the mandatory policy-making process in this chapter versus that described by Dawson and Segerson (2008). In the latter, the regulator's objective is to achieve some aggregate emission level, and in our model, it is to minimize social costs. Because some firms might not join the VP, the mandatory standard is better than the VP for achieving some aggregate abatement level. Without political difficulties, the regulator can implement the mandatory policy, which achieves a socially-efficient outcome. Thus, the difference between this chapter and Dawson and Segerson's paper is crucial to the implementation of a VP by the government.

Furthermore, in our model, the mandatory policy is implemented if the polluting firms think that few firms will participate in the VP. Firms are likely to believe that enough firms will participate if they feel obliged to prioritize environmental issues in corporate-social responsibility (CSR). In the mid-1980s, environmental activists started to increase pressure on firms to manage environmental issues in the US. Consequently, the industry undertook environmental initiatives. The 33/50 program, mentioned above, was implemented in the early 1990s, when many firms began to consider environmental issues as their social responsibility[8]. Therefore, the existence of situations of equilibrium related to the mandatory policy seems consistent with the fact that few voluntary approaches were implemented prior to the 1990s.

The VP is more likely to be implemented if firms feel that they have to tackle environmental issues as part of their CSR. Firms think that participation in the VP appeals to the regulator, even when the VP is not implemented. Therefore, firms gain a small positive benefit (smaller than the abatement cost) from their commitment to join the VP. However, the regulator does not recognize this benefit. As long as the participation rate is low enough that the regulator does not implement the VP (as long as (2.10) holds), ,firms prefer to join because the regulator sets the reduction rate (slightly) lower than α^V. Therefore, firms get the benefit of joining the VP. Thus, a sufficient number of firms always join the VP.

Finally, it is worth mentioning that there is a difference in the number of participating polluters between voluntary programs and international environmental agreements (IEA). Both are collective actions intended to reduce emissions, and polluters can choose whether or not to participate. The number of VP-participating polluters is quite different from that of IEA-participating polluters (when the abatement cost function is linear and the damage function is quadratic). On one hand, Proposition 4 of Barrett (1994) implies that the number of IEA-participating polluters is always fewer than two. On the other hand, all or most polluters might participate in a VP, although the number of participating polluters depends on particular parameters, especially the influence of political contributions on legislators. This difference is due to the existence of a punishment for non-participation for all polluters: mandatory regulation. The strength of this punishment depends on how much the legislator is influenced by political contributions.

---

[8] According to survey data, about 50 percent of American companies had a formal environmental policy statement or added environmental responsibility to company ethics statements by 1992 (Berenbeim (1992)).

2.4 Comparative static analysis of an equilibrium where the most effective VP is implemented

In this section, we examine how changes in parameters affect the abatement rate of individual participating firms, the participation rates, and the aggregate abatement of the most effective VP. For example, if an industry is more pollution-intensive, what happens to the participation rate, the reduction rate, and the aggregate abatement under a VP? Relative to the mandatory policy, how are the reduction rate and the aggregate abatement affected by the pollution-intensity of the industry? What is the role of industry size, marginal costs, marginal damages, and political difficulty in setting an effective mandatory policy? The following proposition shows the effects of changes in these variables on the participation rate, the abatement rates of individual participating firms, and the aggregate abatement level. First, we examine the effects of parameter changes on the number of participating firms, abatement rates and aggregate abatement under the VP.

Proposition 2.3. *If the industry is more pollution-intensive, the size of the industry (the number of the polluting firms) is larger, or the slope of marginal damage is steeper, then the participation rate, reduction rate, and aggregate reduction under a VP increase.* ($\frac{\partial(N_P^V/N)}{\partial\bar{e}_i} \geq 0$, $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial\bar{e}_i} \geq 0$, $\frac{\partial\alpha^V}{\partial\bar{e}_i} \geq 0$, $\frac{\partial(N_P^V/N)}{\partial N} \geq 0$, $\frac{\partial\alpha^V}{\partial N} \geq 0$, $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial N} \geq 0$, $\frac{\partial(N_P^V/N)}{\partial d} \geq 0$, $\frac{\partial\alpha^V}{\partial d} \geq 0$, *and* $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial d} \geq 0$.)

Proof: See Appendix A.3.

This result is quite intuitive (except for N). When the industry is pollution-intensive, the size of the industry is large, or if the slope of the marginal damage is steep, the socially optimal abatement rate and aggregate abatement are large. In such cases, the abatement rate and the aggregate abatement under the mandatory policy must also be high. Therefore, Proposition 3 implies that if the mandatory policy is more stringent, then the VP is also more stringent in the sense that the abatement rate, participation rate and aggregate abatement are higher. Relationships between pollution intensity and the participation rate in the proposition are consistent with the empirical findings of the 33/50 program and the Canadian ARET program(e.g. Antweiler and Harrison (2007) and Gamper-Rabindran (2006)).

The impact of the increase in the size of the industry (the number of firms) on the participation rate may not be intuitive because the larger the size of industry, the stronger the incentive firms have to free-ride. However, lobby activity and the VP suffer from free-riding. In addition, the socially optimal aggregate emission level does not change even if the number of firms increases[9]. Because the socially optimal aggregate emission level does not change, the socially optimal aggregate abatement increases as much as the aggregate emission level if the number of firms increases. Therefore, the socially optimal abatement rate of individual firms increases as the number of firms increases. Due to these two factors, the mandatory standard, a kind of punishment for non-cooperation for the VP, is much more stringent under a larger industry, and the VP is also effective.

Because an increase in the number of firms (and emissions per firm and slope of marginal damage) makes increased regulations desirable, it also makes the actual regulation level (abatement rate of the mandatory standard and VP and participation rate of the VP) higher. This is Proposition 2.3. It seems natural that the actual regulation level is higher if the desirable regulation level is higher. However, if the desirable regulation level is the same, how does the number of firms influence the VP? Intuitively, due to free-riding, the VP for the small number of large polluting firms is more effective than that for the large number of small firms in such a case. However, the following Proposition shows that this intuition is wrong due to the free-riding on the VP.

Proposition 2.4. *The VP under a larger industry is more effective than under a smaller industry if the aggregate natural emission levels are the same (and if the abatement cost, slope of marginal damage and political difficulty in setting the mandatory policy are the same).(If* $N\bar{e}_i = N'\bar{e}'_i$, $N > N'$, $c = c'$, $d = d'$, *and* $\lambda = \lambda'$, *then* $\alpha^V > \alpha^{V\prime}$ *and* $N_P^V/N > N_P^{V\prime}/N'$ *and* $N_P^V\alpha^V\bar{e}_i > N_P^{V\prime}\alpha^{V\prime}\bar{e}'_i$*.)*

Proof: See Appendix A.4.

If the aggregate natural emission levels and other parameters (except for the pollution intensity of individual firms and the size of the industry) are the same, then the socially optimal aggregate abatement levels and rates are also the same. Proposition.4 implies that the VP for the larger industry is more effective when the socially optimal aggregate abatement levels and rates are the

---

[9] This is because the marginal aggregate abatement cost is constant and independent of the number of firms.

same. This result is counterintuitive, but it indicates that the effectiveness of lobbying is eroded by free-riding. Remember $\alpha^L = 1 - \frac{c}{dN\bar{e}} - \frac{c\lambda}{dN^2\bar{e}(1-\lambda)}$. The third term reflects the effect of lobbying on the level of the mandatory standard, and it is greatly affected by the size of the industry because the third term is inversely proportional to N square. Thus, the industry size has significant effects on the effectiveness of lobbying when lobbying suffers from free-riding.

The effectiveness of lobbying directly determines the abatement rate of the VP and indirectly determines the participation rate of the VP or the seriousness of free-riding on the VP. First, we explain a mechanism that determines the participation rate, then we explain how the mechanism works when lobbying is ineffective. The participation rate depends on the difference between the abatement rates of individual firms under the VP and the mandatory policy. If this difference is large (individual firms abate greatly under the VP relative to the mandatory policy), the participation rate is low because the VP can generate a higher aggregate abatement than the mandatory policy even though the participation rate is low. If not, only a VP with a high participation rate can generate higher social welfare than the mandatory policy. Thus, the difference between the abatement rates determines the participation rate of the VP or the degree of free-riding.

The difference between the abatement rates of the VP and the mandatory policy is determined by the resources of individual firms reallocated from lobbying efforts to abatement under the VP. If the lobbying efforts are low, the difference between their abatement rates is small, and therefore, the participation rate is high according to the discussion in the last paragraph. Remember that the lobbying effort is small due to free-riding when the industry is large. Thus, we obtain results similar to Proposition 2.4[10].

We also examine the effects of change in parameters that lower the socially optimal abatement rate or the abatement rate under the mandatory standard. As we can guess from Proposition 3, the VP is also less stringent if the mandatory policy is less stringent.

---

[10] However, if lobbying does not suffer from free-riding, the participation rate decreases as the number of firms increases provided that aggregate emissions stay the same. In Appendix A.5, we analyze a case where no firms free-ride on the lobbying of others. We formally show that under the assumption of Proposition 2.4, the VP is effective and the participation rate is high if the industry is small (Proposition A.2). In constrast with Proposition 4, Proposition 3 and the propositions shown below (Propositions 2.5 and 2.6) hold even though the lobbying activity does not suffer from free-riding.

Proposition 2.5. *If the industry's marginal costs increase or if the legislator is influenced heavily by political contributions, then the participation rate, reduction rate, and aggregate reduction under a VP decrease. (*$\frac{\partial(N_P^V/N)}{\partial c} \leq 0$*,* $\frac{\partial \alpha^V}{\partial c} \leq 0$*,* $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial c} \leq 0$*,* $\frac{\partial(N_P^V/N)}{\partial \lambda} \leq 0$*,* $\frac{\partial \alpha^V}{\partial \lambda} \leq 0$*, and* $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial \lambda} \leq 0$*.)*

Proof: See Appendix A.3.

Proposition 2.5 implies that the participation rate is likely to be low if λ is high. A high λ can be interpreted as high political difficulty in setting the mandatory policy. Therefore, it may be good to implement the voluntary program with a low participation rate, as occurred with the 33/50 program or the ARET program, if it is politically difficult to set the mandatory standard when the program is initiated.

The above two propositions explain the impact of changes in parameters on the VP but do not state that the impacts on the VP are smaller or greater than they are on the mandatory policy. By comparing the impacts on the VP with those on the mandatory policy, we can evaluate their magnitude and better characterize VPs. The next proposition describes the impact on VPs relative to that on the mandatory policy.

Proposition 2.6. *If a parameter changes, the abatement rate of the VP will change less than under a mandatory policy, but the aggregate abatement of the VP will change more. For* $\bar{e}_i$*,* $0 \leq \frac{\partial \alpha^V}{\partial \bar{e}_i} \leq \frac{\partial \alpha^L}{\partial \bar{e}_i}$*, and* $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial \bar{e}_i} \geq \frac{\partial N \alpha^L \bar{e}_i}{\partial \bar{e}_i} \geq 0$ *(The same relationships hold for N, and d). For c,* $0 \geq \frac{\partial \alpha^V}{\partial c} \geq \frac{\partial \alpha^L}{\partial c}$ *and* $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial c} \leq \frac{\partial N \alpha^L \bar{e}_i}{\partial c} \leq 0$ *(The same relationships hold for* λ*).*

See Appendix A.3.

Proposition 2.6 gives us two results. First, we explain why the reduction rate of the VP increases less than that of the mandatory policy. Changes in parameters affect political contributions. If political contributions decrease, then α^L increases. Although the total cost of firms under a mandatory policy increases due to an increase in the abatement cost, the increased rate of the total cost is smaller than that of the abatement cost because political contributions decrease. This change in total cost under the mandatory policy changes the abatement rate of the VP less than that of the mandatory policy because the VP's abatement rate depends on the total cost under the mandatory policy.

Proposition 2.6 also states that aggregate abatement under the VP is more sensitive to changes in parameters than it is under the mandatory standard. Changes in parameters affect the aggregate abatement under the VP through two channels: the number of participating firms, N_P^V, and the individual participating firms' abatement rate, α^V. The effect of α^V on the aggregate abatement under the VP is smaller than that of α^L given the aggregate abatement under the mandatory standard. However, due to the effect of N_P^V, changes in parameters affect the aggregate abatement under the VP more than they do under the mandatory standard. More technically, it should be recalled that $N_P^V = N\alpha^L/\alpha^V + 1$, and therefore the aggregate abatement under the VP is a sum of the aggregate abatement under the mandatory standard and one firm's abatement under the VP ($N_P^V \alpha^V \bar{e}_i = N\alpha^L \bar{e}_i + \alpha^V \bar{e}_i$). The first term is included by (2.10), the constraints on the number of participating firms[11], and the second term is determined by (2.8), the constraints on the abatement cost or the abatement rate. We can divide the impact of the changes in parameters on aggregate abatement into those that affect the conditions on the number of participating firms and those that affect the abatement rate. Because the impact on the former condition is the same as the impact on the aggregate abatement under the mandatory policy and the aggregate abatement under the VP is also influenced by the latter, changes in parameters influence aggregate abatement under the VP more than they do under the mandatory policy.

2.5 Conclusion

We build a model with a regulator, a legislator and multiple polluting firms in which the legislator, who is affected by a lobby group representing polluting firms, sets the mandatory standard. We then used the model to explain why the regulator implements a VP. In this model, the regulator can implement the VP, which generates lower social costs than the mandatory standard, in contrast to

---

[11] (2.10) is not the direct constraint on the number of participating firms but rather that on the aggregate abatement. However, only the number of participating firms is determined by (2.10) because the abatement rate is determined by (2.8).

the model of Dawson and Segerson (2008). This difference in models occurs because we introduced an element of political economy into a mandatory policy-making process and the regulator in their model had a different objective from the regulator in this chapter's model. The regulator's objective in Dawson and Segerson (2008) was to achieve some aggregate emission level, but it was to minimize social costs in this chapter. Because the VP is subject to free-riding, the mandatory standard better achieves an aggregate abatement level than the VP does. Without political influence, the regulator could implement the mandatory policy, which would achieve a socially-efficient outcome. Thus, the differences between this chapter's model and Dawson and Segerson's shed important light on the ways a government should implement VPs.

We found that the regulator should set the abatement rate under the VP at the highest possible level and should not set the abatement rate to maximize the participation rate. However, setting the reduction rate of the VP such that all polluting firms participate in the VP might be optimal when the legislator responds weakly or not at all to the lobbying of the polluting industries. Otherwise, setting the reduction rate at this level is not optimal.

This study assumed that all firms are identical. However, one way to extend this chapter's model is to introduce the heterogeneity of polluting firms. Given that most voluntary approaches cannot enforce a firm's commitment, it would be interesting to consider the case in which voluntary programs are not enforceable. Legislators have the right to make laws, and it is thus possible that legislators would set a mandatory standard even if a government or environmental organization decides to implement a VP. Exploring these types of extension remains an endeavor for future research.

Chapter 3

Intra-Industry Spillover Effects of ISO 14001 Adoption and Environmental Performance in Japan

3.1 Introduction

Voluntary approaches have become increasingly popular approaches to environmental challenges. One of the reasons for the popularity of voluntary approaches is that they are much more flexible than traditional command-and-control interventions. Voluntary approaches allow firms to reduce emissions through more cost-effective methods than the often less effective mandated methods. Another reason for the popularity of voluntary approaches is that incentive-based mechanisms impose additional costs on firms beyond the expenses of pollution abatement, such as emission taxes or emission permit purchases. For these reasons, voluntary approaches are more acceptable to firms than command-and-control or incentive-based mechanisms. Governments are promoting voluntary actions to address environmental issues for which it is difficult to employ mandatory policies.

The introduction of environmental management systems (EMSs) is one of the most common voluntary actions performed by firms. In particular, the ISO 14001 standard has received increasing attention. According to the International Organization for Standardization (ISO), a total of 223,149 certificates had been issued worldwide by the end of 2009. In response to the popularity of ISO 14001, researchers have examined the reasons why facilities adopt this standard. For example, by analyzing Japanese company data, Nakamura et al. (2001) and Nishitani (2009) find that certain characteristics of firms, such as their size, export ratio, debt ratio, pressures from stakeholders, and their financial flexibility, affect their ISO 14001 adoption. Others also find that firms are certified to ISO 14001 earlier under greater regulatory pressures (King et al. (2005), Potoski and Prakash (2005a,b), and Darnall and Edwards (2006)).

However, ISO 14001 has been criticized because it does not explicitly specify an objective or target for environmental performance. ISO 14001 is focused on operational processes but not environmental outcomes. To obtain ISO 14001 certification, firms have to establish a Plan-Do-Check-Action (PDCA) cycle, which is a continual cycle of planning (plan), implementing (Do), reviewing and improving the processes (Check) and actions (Action) that are aimed at meeting the firms' own environmental targets and continually improving their environmental performance. Although ISO 14001 does not include explicit environmental performance objectives, it may nevertheless help firms improve their environmental performance. Therefore, the effectiveness of ISO 14001 adoption has also been examined in many studies. Some studies find little evidence that ISO 14001 adoption has a positive effect on environmental performance (Barla (2007), King et al. (2005), and Darnall and Side (2008)), whereas a number of other studies find substantial evidence of that effect (Arimura et al. (2008), Potoski and Prakash (2005a,b), and Melnyk et al. (2003)). A key contribution of this chapter is to examine whether ISO 14001 has made a positive contribution to the environmental performance of Japanese firms.

Spillovers between firms were not considered in previous studies on the adoption and effectiveness of ISO 14001, except for Arimura et al. (2009), who estimates the effects of ISO adoption on green supply chain management. If spillover effects exist among firms within the same industry, estimation results of the effect of ISO 14001 on environmental performance will be biased. According to an international survey conducted by the Organization for Economic Co-operation and Development (OECD) in Canada, France, Germany, Hungary, Japan, Norway, and the US, the fact that similar facilities are adopting EMS and similar environmental practices motivates other firms to adopt them. Thus, the decisions by firms and facilities on ISO 14001 adoption and environmental performance are likely to be influenced by those of other firms and facilities within the same industry. If such externalities exist, it is important not only to control for them but also to measure their magnitudes when we estimate the effect of contributing factors on ISO 14001 adoption and environmental performance. In addition, if intra-industry spillovers exist, governments can pursue industry-specific voluntary programs in cooperation with industry associations.

Employing an extensive dataset of Japanese production facilities across a large number of industries, this chapter examines the determinants of ISO 14001 adoption, the effect of ISO 14001 adoption on environmental performance, and the existence of intra-industry spillovers. To control for and estimate the spillovers, we employ a Bayesian spatial autoregressive (SAR) probit model for ISO 14001 adoption and a SAR model for two types of emission reductions (emissions into air and emissions into water). The potential for intra-industry spillovers is based on the hypothesis that facilities that adopt ISO 14001 (or reduce emissions) do so because other facilities in the same industry also adopt ISO 14001 (or reduce emissions). This type of spillover has been examined by a few empirical studies on voluntary environmental approaches (including ISO 14001). Only Antweiler and Harrison (2007) considers spillovers, and their results show an intra-industry spillover of participation in the Canadian voluntary ARET (Accelerated Reduction/Elimination of Toxics) program. However, they do not estimate the effect of spillovers on emission reductions.

By employing spatial econometric estimation methods, we find that there are positive spillovers of ISO 14001 adoption for reducing water emissions between Japanese facilities that belong to firms with similar revenue levels in the same industry. Plants that adopt ISO 14001 are more likely to do so if their industry peers also adopt ISO 14001. In addition, the percentage change in the weighted sum of emissions into the air, used as a measure of environmental performance, is correlated between similar sized facilities in the same industry. This correlation might reflect the fact that similar sized facilities in the same industry have similar technologies and therefore have similar environmental performance.

There are three other noteworthy results. First, plant size (as determined by the number of workers in a facility) has a significantly positive effect on ISO 14001 adoption, confirming the results from previous studies on ISO 14001 adoption. Second, facilities that emit into water with lower water emission intensities compared to other facilities are more likely to adopt ISO 14001. Third, there is no statistically significant evidence that ISO 14001 adoption improves environmental performance.

This chapter is organized as follows. Section 2 briefly explains the background of this research, the ISO 14001 standard, and the reasons for employing a Japanese dataset for this study. This section also discusses the rationale for examining intra-industry spillovers. We explain our econometric models in Section 3, and in Section 4, we introduce the data from the empirical analysis of this chapter. Section 5 shows estimation results, and then we conclude the paper in Section 6.

3.2 Background and hypotheses on spillover effects

3.2.1 ISO 14001 in Japan

In this section, we explain the reasons for employing a Japanese dataset to examine the determinants of ISO 14001 adoption and its effect on environmental performance and why we have to examine intra-industry spillover effects.

ISO 14001 is an internationally recognized standard for an environmental management system (EMS) that was released in 1996 and revised in 2004 by the International Organization for Standardization. Firms can adopt ISO 14001 at the level of the individual facility, group(s) of facilities, or the entire company. To adopt ISO 14001, facilities are certified by external third-party registrars. Certified facilities must follow a cycle of Plan-Do-Check-Act over time: environmental planning ("Plan"), plan implementation and operation ("Do"), monitoring ("Check"), corrective action ("Action"), and management review.

As shown in Table 3.1, ISO 14001 has been very popular in Japan since its release. ISO 14001 adoption has increased more rapidly in Japan than in other countries except China after 2003. Actually, Japan's share of the total number of ISO 14001 certificates was more than 20% until 2005. Although the total number of certificates in China (39,195) exceeded that of Japan (35,573) in 2008, this number is more than twice that of Spain, which had the third highest number of certified organizations (16,443). Due to the popularity of ISO 14001 in Japan, many studies on ISO 14001 employ Japanese data at the firm or facility level (e. g. , Arimura et al. (2008), Arimura et al. (2009), Nakamura et al. (2001), Nishitani (2009), and Welch et al. (2002)). The high adoption rate of ISO 14001 across a large number of industries in Japan is a prime reason for employing such data in this study as well.

Table 3.1: Top 3 for ISO 14001 certificates in 2000, 2003, and 2008

2000		2003		2008	
Japan	5556	Japan	13416	China	39195
UK	2534	UK	5460	Japan	35573
Sweden	1370	China	5064	Italy	16443

Previous studies of Japanese facilities or firms found that facility/firm size has a significant positive impact on the adoption of ISO 14001 environmental management standards (see Arimura et al. (2008), Arimura et al. (2009), and Welch et al. (2002) for facility level analysis and Nakamura et al. (2001) and Nishitani (2009) for firm level analysis). This relationship likely exists because the costs of ISO 14001 adoption are less significant for large facilities or firms than for small ones. In addition, foreign customers who find it difficult to monitor the performance of overseas firms may require ISO 14001 adoption as a visible sign of commitment to environmental protection. Therefore, facilities or firms with more foreign customers are more likely to adopt ISO 14001; previous studies have found evidence to confirm this conclusion (Arimura et al. (2009), Arimura et al. (2008), Nakamura et al. (2001) and Nishitani (2009)).

ISO 14001 adoption in Japan might be stimulated by governmental policies. Many local governments have encouraged the adoption of ISO 14001 through financial support and/or informational support, such as seminars on ISO 14001. Facilities in our sample were located in 1,056 municipalities, and 105 of them provided some support for ISO 14001 adoption in 2001. Some local governments have also adopted ISO 14001 and have provided advice on ISO 14001 adoption based on their own experience.

However, in recent years, Japanese governments have encouraged EMS certifications other than ISO 14001 because most large firms adopted ISO 14001 around 2005, while it remained too expensive for small and medium-sized enterprises (SMEs) to adopt ISO 14001. Therefore, Japanese governments started encouraging SMEs to adopt EMSs that are less expensive than ISO 14001, such as Eco-Action 21, which was launched by the Japanese Ministry of the Environment in 1996 and became an EMS with a third-party certification in 2004. Because complete information is not available on the registration of certain EMSs that target SMEs, we employ emission data during

2001-2003, when there was a small number of registrations of these alternative EMSs.

The EMS adoption decision by facilities likely influences the EMS adoption decisions of other firms in the same industry. An OECD survey, “Environmental Policy Tools and Firm-Level Management and Practices in Japan, ” shows that 53% of Japanese facilities consider what “other facilities like ours are adopting” as “very important” or “important” as motivation for adopting an EMS(Hibiki and Arimura (2004)). Therefore, it is very likely that spillover effects of ISO 14001 adoption exist within an industry. The above OECD survey was conducted not only in Japan but also in Canada, France, Germany, Hungary, Norway, and the US. This multi-country survey also asked firms about their motivations to implement environmental practices. Specifically, firms were asked about the importance of the fact that similar facilities were adopting similar environmental management practices[1]. From 20% to 60% of firms consider that this fact is moderately important or very important (Darnall and Pavlichev (2004), Glachant et al. (2004), Kerekes et al. (2004) Rennings et al. (2004), and Ytterhus (2004)). It is very likely that some, or even many, Japanese firms think and act similarly. Therefore, both the decision to adopt ISO 14001 and the decision to improve environmental performance may be influenced by the actions of other firms (or their facilities).

3.2.2 Hypotheses on spillovers

How should we construct weight matrices to capture spillover effects? The answer depends on what types of spillover effects we want to examine. In this subsection, we discuss hypotheses about the effects of intra-industry spillover on the decision to adopt ISO 14001 and on environmental performance.

As mentioned in the previous subsection, the OECD survey revealed that many Japanese facilities were motivated to adopt ISO 14001 by the knowledge that other facilities/firms in the same industry had adopted the standards. If we interpret this evaluation literally, industry-wide spillover effects likely exist for the adoption of ISO 14001: facilities adopt ISO 14001 because other facilities in the same industry have adopted it. In addition, the OECD survey showed that many facilities

---

[1] This question was posed to firms in Canada, Germany, Hungary, Norway, and the US but unfortunately not to firms in Japan.

considered the environmental practices adopted by similar facilities to be an important motivation for the adoption of similar practices. If facilities with similar levels of capital intensity and technology adopt similar environmental practices, their environmental performance is likely correlated and may reflect intra-industry correlation or spillover effects. Therefore, we examine three main hypotheses.

Hypothesis 1 There are industry-wide spillover effects or correlations that affect ISO 14001 adoption and environmental performance.

As discussed above, the OECD survey indicates that facilities may imitate or adopt the same practices as other facilities. Such imitation can occur for competitive reasons. For example, firms are likely to adopt the same practice if they think that not doing so would reduce their competitiveness in the marketplace. Such imitation pressure for competitive reasons is likely to be strong when the practice is widespread. The adoption of ISO 14001 was already widespread in Japan in 2000, as seen in Table 3.1. Imitation pressure because of competition may therefore contribute to the diffusion of ISO 14001.

A facility/firm is likely to have a greater influence on the decisions and actions of similar facilities/firms than on those with different characteristics. This is especially true if the most significant determinants of the decisions and actions of facilities differ; those facilities will have little influence on each other and will act more independently. For example, many studies, such as Nakamura et al. (2001), Welch et al. (2002), Arimura et al. (2008) and Nishitani (2009), have found that facilities with more workers are more likely to adopt ISO 14001; this is likely true because of the comparatively high initial cost of adoption. This finding implies that large facilities have a strong incentive to adopt ISO 14001, whereas smaller facilities may not have an incentive to adopt it. The adoption of ISO 14001 by large facilities may not affect the decisions of small facilities. There may also be a negative correlation between ISO 14001 adoption at large and small facilities because they are likely to have very different adoption incentives.

Therefore, we consider spillover effects and correlations among similar facilities. In particular, we focus on two types of similarities. First, we focus on facility size as measured by the number of workers. As mentioned above, facility size is a significant determinant of ISO 14001 adoption. This makes logical sense because facilities that have more workers are likely to be able to allocate more workers to ISO 14001 adoption. Thus, facilities with a similar number of workers may have similar attitudes toward ISO 14001 adoption and therefore make similar decisions. Thus, intra-industry correlations or spillover effects may exist for ISO 14001 adoption at similarly sized facilities.

In addition, there may be intra-industry correlations or spillover effects for environmental performance at similarly sized facilities. Similarly sized facilities in the same industry are likely to have similar levels of capital intensity and technologies, and they are therefore likely to have similar environmental performance.

Hypothesis 2 There are intra-industry spillover effects or correlations between similarly sized facilities that affect ISO 14001 adoption and environmental performance.

The second type of similarity we focus on is the size of the firms to which facilities belong. Similarly sized firms in the same industry are likely to be competitors and therefore may adopt similar environmental practices. Among the variables in our dataset, revenue is most able to capture firm size or demonstrate “rivalry". Therefore, facilities that belong to firms with similar levels of revenue may influence each other’s decisions and actions on environmental issues.

Hypothesis 3 There are intra-industry spillover effects or correlations among facilities that belong to similarly sized firms (as determined by revenue) that affect ISO 14001 adoption and environmental performance.

Intra-industry spillover effects among similarly sized firms (as determined by revenue) is likely to explain the ISO 14001 adoption decisions of facilities better than the spillover effects among similarly sized facilities. Because of the initial cost, the ISO 14001 adoption decisions of many facilities are not made by the facilities themselves but by the firms to which they belong. However, spillover effects among similarly sized facilities are likely to explain environmental performance better than the other hypotheses because facility size may be a more important determinant of environmental performance. In the next section, we will discuss our strategy for evaluating these three hypotheses.

3.3 Estimation strategy

The models that we employ are essentially spatial econometric models. For emission reduction, we employ a spatial autoregressive (SAR) model. For ISO 14001 adoption, we employ a SAR probit model. We will not discuss the estimation methods in detail here; LeSage and Pace (2009) for details (in particular, Chapters 3, 5, and 10).

Let E_{ijklt} and $\Delta E_{ijklt}(= E_{ijklt+1}/E_{ijklt})$ be the emission level of facility i of company j in industry k at municipality l and the emission reduction (change) of facility i from year t to year t+1, respectively. We normalize both E_{ijklt} by facility size (the number of workers in the facility) and denote by EL_{ijklt} this normalized emission level. We can interpret EL_{ijklt} as the proxy of emission intensity in terms of facility size because the facility size is very likely to be closely correlated with the number of workers. We employ EL_{ijklt} as the proxy of emission intensity due to data availability issues although it might be better to define emission intensity dividing by output or revenue. We will discuss data availability in the next section.

We assume that the emission reduction equations take the following form.

$$\Delta E_{ijklt} = \rho \sum_m w_{im} \Delta E_{mjklt} + \beta_F X_{it}^F + \beta_C X_{jt}^C + \beta_M X_{lt}^M + \gamma I\hat{S}O_{ijklt} + \varepsilon_{ijklt} \tag{3.1}$$

or

$$\Delta E_t = \rho W \Delta E_t + \beta_F X_t^F + \beta_C X_t^C + \beta_M X_t^M + \gamma I\hat{S}O_t + \varepsilon_t. \tag{3.2}$$

X_{it}^F, X_{jt}^C, and X_{lt}^M are characteristics of facility i, firm j, and municipality l, which are explained below. W is a (standardized) weight matrix whose elements are $w_{ij} = 1/m_i$ if facility j is one of the m_i nearest neighbors of i (but $w_{ii} = 0$), and $w_{ij} = 0$ otherwise. We will discuss the selection criteria for the nearest neighbors at the end of this section. To control for the endogeneity of ISO 14001 adoption, $I\hat{S}O_{ijklt}$ is the ISO 14001 adoption modeled using a latent variable, ISO_{ijklt}^*. Concretely, $I\hat{S}O_{ijklt} = 1$ if $ISO_{ijklt}^* > 0$ and $I\hat{S}O_{ijklt} = 0$ otherwise. The ISO 14001 adoption latent variables are estimated by

$$\begin{aligned} ISO_t^* &= \nu WISO_t^* + \kappa_F X_t^F + \kappa_C X_t^C + \kappa_M X_t^M + \lambda EL_t + \mu_t \qquad (3.3) \\ \mu_t &\sim N(0, I_n). \end{aligned}$$

In the next subsections, we briefly explain our methods and procedures for estimating (3.2) and (3.3).

3.3.1 Estimation of environmental performance equation

We estimate (3.2) by Maximum likelihood (ML). Concretely, we assume $\varepsilon_t \sim N(0, I_n)$ and minimize the following log-likelihood function

$$\min_{-1<\rho<1} \log L(\rho) = -(n/2)\log(\pi) + \log \mid I_n - \rho W \mid -(n/2)\log(S(\rho))$$

where $S(\rho) = (e_o - \rho e_d)'(e_o - \rho e_d)$, $e_o = y - X\hat{\beta}_o$, $e_d = y - X\hat{\beta}_d$, $\hat{\beta}_o = (X'X)^{-1}X'y$,$\hat{\beta}_d = (X'X)^{-1}X'Wy$, and $X = (X^F, X^C, X^M)$. Note that $-1 < \rho < 1$ must hold for the minimization problem to be solvable. After solving this optimization problem and determining optimum $\hat{\rho}$, we then derive $\hat{\beta} = (\hat{\beta}_F', \hat{\beta}_C', \hat{\beta}_M')'$ from $\hat{\beta} = \hat{\beta}_o - \rho\hat{\beta}_e$.

3.3.2 Estimation of the ISO 14001 adoption equation

We estimate the equation (3.3) by MCMC sampling which is a standard method for estimating spatial probit models. We need conditional distributions in order to sample the latent variable ISO_{ijklt}^* and need prior distribution(s) $(\nu, \kappa_F, \kappa_C, \kappa_M, \lambda)$ to derive the conditional contributions. We assume prior independence between ν, and θ, where $\theta = (\kappa_F, \kappa_C, \kappa_M, \lambda)$. We also assume a univariate uniform pdf $U(-1, 1)$ for ν, and a multivariate-normal pdf $N(c, T)$ for θ.

Let the observed ISO 14001 adoption, ISO_{it}, be equal to 1 if facility i adopts ISO 14001 at year t and equal to 0 otherwise. Then, the joint posterior distribution is given by

$$\pi(\theta, \nu|ISO_{it}) \propto |A|exp(\frac{1}{2}\mu_t'\mu_t)\pi(\theta)\pi(\nu) \tag{3.4}$$

where $A = I_n - \nu W$ and $\pi(\theta)$ and $\pi(\nu)$ are prior on θ and ν. The conditional distributions are

$$\pi(\theta|\nu, ISO_t, ISO_t^*) \propto exp((AISO_t^* - X\theta)'(AISO_t^* - X\theta))exp((\theta - c)'T^{-1}(\theta - c)) \tag{3.5}$$

$$\pi(\nu|\theta, ISO_t, ISO_t^*) \propto exp((AISO_t^* - X\theta)'(AISO_t^* - X\theta)) \tag{3.6}$$

$$\pi(ISO_t^*|\theta, \nu, ISO_t) \propto exp((AISO_t^* - X\theta)'(AISO_t^* - X\theta)) \tag{3.7}$$

where $X_t = (X_t^F\ X_t^C\ X_t^M\ E_t)$. (3.5) implies that θ is distributed f^{MVN} with mean M and variance V where $M = (X'X + T^{-1})^{-1}(X'AISO_t^* + T^{-1}c)$ and $V = (X'X + T^{-1})^{-1}$.

We sample the ν using a random-walk Metropolis-Hastings procedure because they do not correspond to any known probability distribution. Candidate variables ν^C are produced by using standard normal distributions and tuning parameters d as follows:

$$\nu^C = \nu^* + dN(0,1)$$

Then, we accept these candidates by probability,

$$p = \min\left[1, \frac{f(\nu^C|\theta, ISO_t, ISO_t^*)}{f(\nu^*|\theta, ISO_t, ISO_t^*)}\right], \tag{3.8}$$

where

$$f(\nu^*|\theta, ISO_t, ISO_t^*) = ((AISO_t^* - X\theta)'(AISO_t^* - X\theta)). \tag{3.9}$$

By following LeSage and Pace (2009), we adjust tuning parameters based on monitoring the acceptance rates from the Metropolis-Hastings procedure during the MCMC sampling. If the acceptance rate is smaller than 40%, then we update $d' = d/1.1$. If the acceptance rate is greater than 60%, then $d' = 1.1d$.

(3.7) is proportional to the multivariate normal distribution with mean $A^{-1}X\theta(=\kappa=(\kappa_1,\kappa_2,\cdots,\kappa_n)')$and variance $(A'A)^{-1}$ subject to truncation constraints, which depends on the observed value 0 or 1 for ISO. We sample the individual element ISO^*_{it} using Geweke's (1991) approach. We sample z_i conditional on z_{-i} under the truncation constraints and construct $ISO^*_{it}=\kappa_i+z_i$. Let $\Psi=A'A$and $\gamma_{-i}=\Psi_{i,-i}/\Psi_{ii}$ where $\Psi_{i,-i}$ is the$i$th row of Ψ without the ith element. Then, $E(z_i|z_{-i})=\gamma_{-i}z_{-i}$, and therefore, a normal conditional distribution for z_i is obtained as follows;

$$z_i|z_{-i}=\gamma_{-i}z_{-i}+(\Psi_{ii})^{-1/2}v_i \tag{3.10}$$

where $v_i\sim N(0,1)$ under the truncation constraints,

$$\begin{aligned} v_i &< (-\kappa_i-\gamma_{-i}z_{-i})(\Psi_{ii})^{1/2} \quad (\Leftrightarrow ISO^*_i<0) \quad if\, ISO_i=0 \\ v_i &> (-\kappa_i-\gamma_{-i}z_{-i})(\Psi_{ii})^{1/2} \quad (\Leftrightarrow ISO^*_i>0) \quad if\, ISO_i=1. \end{aligned}$$

We sample v_i from $N(0,1)$ under the truncation constraints. We implement m-step Gibbs sampling to produce z. On the initial step, we set z to zeros, and at the jth iteration of z_i, we sample z_i^j using $z_1^j,z_2^j,...,z_{i-1}^j$and $z_{i+1}^{j-1},z_{i+2}^{j-1},...,z_n^{j-1}$ where $z_i^j$is the value after the j iteration. We continue this procedure for m iterations for all i.

Our sampling routine for estimating the parameters is

1. determine θ by drawing from $f^{MVN}(M,V)$;
2. determine ν by generating candidate ν using a random-walk procedure and accepting them with probability p in (3.8);
3. determine ISO^* by generating z^m by sampling v_i and updating z_i^j for $i=1,2,..,n$ and $j=1,2,...,m$ as explained above, and setting $ISO^*=\kappa+z^m$.

3.3.3 Weight matrices

We employ three types of weight matrices to capture and examine the intra-industry spillover effects discussed in Section 2. The first weight matrix is intended to capture and examine spillover effects between facilities in the same industry. Because facilities are likely to be more influenced by other facilities with similar characteristics, the second and third matrices are formed to capture and examine spillover effects between "similar" facilities within industries, in terms of the number of workers and revenue levels of the firms. For comparison, we also estimate models with/without spatial correlations with all other facilities. More formally, we estimate models with the following five weight matrices.

1. Weight matrix $W = \mathbf{0}$ (no spatial correlation) (Weight Matrix I (WM I))
2. For each i, $W_{ii} = 0$ and $W_{ij} = 1/(n-1)\ \forall j \neq i$. (correlated with all other facilities) (WM II)
3. For each i, $W_{ij} = 1/m_i\ \forall j \neq i$ if facility $i$and j are in the same industry. (correlated with all other facilities in the same industry, m_i is the number of facilities in the same industry as i) (WM III)
4. For each i, $W_{ij} = 1/m$ for any facility j ($j \neq i$) that is in the same industry as i and one of the m nearest facilities in terms of the number of workers in the facility[2]. (WM IV)
5. For each i, $W_{ij} = 1/m$ for facility j ($j \neq i$) that is in the same industry as i (but belongs to different firms) and one of the m nearest facilities in terms of revenue [3]. (WM V)

For WM IV and V, we employ $m = 3, 5, \text{and} 10$. WM III is constructed to examine Hypothesis 1, WM IV for Hypothesis 2, and WM V for Hypothesis 3.

If some element W_{ij} of W is positive, the decisions or actions of facilities i and j are assumed to correlate with each other or influence each other. Therefore, if W_{ij} is positive, we can examine whether there is a spillover effect or correlation in the decisions and actions of facilities i and j. However, $W_{ij} = 0$ implies that we assume that there is no correlation or spillover effect in the decisions or actions of facilities i and j. Thus, the WM I model is similar to a standard probit model[4] because all elements of W are equal to 0, and there are no correlations between the decisions or actions of various facilities.

---

[2] If several facilities have equal numbers of employees, we consider the facility that belongs to a firm with more similar revenue as nearer.

[3] If several facilities belong to firms with equal revenue, we consider the facility with the more similar number of workers to be nearer.

[4] There is a slight difference because we employ a Bayesian estimation method to estimate this model.

In other models, spillover effects between the decisions or actions of some facilities are thought to exist. In the WM Ⅲ model, we assume that there are correlations between the decisions or actions of all facilities in the same industry but no correlation between any facilities in different industries. Thus, we use this model to examine intra-industry and industry-wide spillovers, i.e., Hypothesis 3.

The nearest facility in terms of number of workers is the facility of most similar size, whereas the nearest facility in terms of revenue is equal to the facility that belongs to the most similarly sized firm (in terms of revenue). Therefore, the WM Ⅳ model captures the intra-industry spillover effects between similar sized facilities, and the WM V captures the intra-industry spillover effects between facilities that belong to similar sized firms. If m is smaller, then we examine spillover effects within narrower limits (between the larger numbers of facilities).

The WM Ⅱ model can capture inter-industry spillover effects. However, we construct this model mainly to check whether the intra-industry spillover effects of the three models above are significant and are not due to unobservable effects on all facilities. In other words, this model is built to examine whether those three models precisely capture the intra-industry spillover effects that we want to examine and whether the intra-industry spillover effects have explanatory power for ISO 14001 adoption and environmental performance.

3.3.4 Model choice

We employ models with 5 types of weight matrices for both the first and second stages. However, we construct $I\hat{S}O_{ijklt}$ by the first stage model which is supported by the Bayesian or Akaike Information Criterion according to Hepple (2004).

3.4 Data description

3.4.1 Emission data

We collect data on the emission and transfer of 354 chemical substances from the PRTR webpage. These data are available from 2001. We consider emissions into air and water and calculate the toxicity-weighted sum of all direct emissions into air and water from each facility ($E^A_{t+1}/E^A_t$and E^W_{t+1}/E^W_t), respectively. For this calculation, we employ the US Environmental Protection Agency (2010) Risk-Screening Environmental Indicators (RSEI) which are also employed by Antweiler and Harrison (2007) and Potoski and Prakash (2005a). We construct the percentage changes in the weighted sum of emissions into air and water (E^A_{t+1}/E^A_t and E^W_{t+1}/E^W_t)and use them as environmental performance indicators because many firms set their emission reduction goals in terms of percentage reduction relative to the previous year.

3.4.2 ISO 14001 adoption data

We can collect data on ISO 14001 adoption and the time of adoption for facilities that were certified by Japanese certification bodies from the Japan Accreditation Board for Conformity Assessment (JAB). Data on ISO 14001 certifications by foreign certification bodies are collected from the firms' websites.

3.4.3 Data on other variables

Data on other characteristics of firms, such as their revenue, profit, and age, are collected from Japanese company data books (Teikoku Data Bank (2003a,b, 2004a,b, 2005a,b)). Pressure from residents near a facility can motivate a company to implement voluntary actions on environmental issues, and the income and population density of the local community are very likely determinants of the level of pressure for a greener environment. Therefore, we collected data on municipal characteristics, such as income, population, area (for calculation of population density) from the CD of the Japanese"System of Social and Demographic Statistics". In addition, the author obtained data on the number of grievances against environmental pollution for each municipality from the Environmental Dispute Coordination Commission to disclose and determined whether the municipalities provided informational and financial support for ISO 14001 adoption. Support from most of the municipalities targeted small and medium-sized enterprises (SMEs), but not all of the SMEs that adopt(ed) ISO 14001 receive(d) the support.

From the PRTR database, we determine the number of employees in each facility and construct the emissions per employee as a measure of emission intensity. To construct the emission intensity, output-based measures, such as revenue or profit, might be more ideal than the number of employees. However, revenue and profit are computed at the firm level rather than at the facility level. In our sample, 252 of 663 facilities that emit into water (38.0%) belong to multi-facility firms, as do 1,793 of 3,579 facilities that emit into air (50.1%). Therefore, firm level variables are not suitable for calculating the emission intensity at the facility level, and therefore, we employ the emissions intensity in terms of workers in the facility. Using employment as a proxy for size might lead to distortions of emission intensity because capital intensity varies across industries. It is reasonable to assume that capital intensities are relatively similar within a given industry. Therefore, emission intensities normalized to employment are comparable within that industry. For comparisons across industries, the industry dummies or fixed effects will account for the differences in capital intensities.

3.4.4 Descriptive statistics

Tables 3.2 and 3.4 show the descriptive statistics of the main variables (in 2001) used for the estimation. Because we estimate the determinants of ISO 14001 adoption and its effect on reduction of emissions into air and water separately, we show the descriptive statistics of facilities that emit into water (Table 3.2) and air (Table 3.4) separately[5]. Note that the average air emission intensity decreased from the year 2001 to 2003, while the average water emission intensity increased during that period. On average, facilities that emit into air were located in municipalities with higher population densities, wealthier residents, and more grievances against environmental pollution compared to facilities that emit into water. In addition, a higher percentage of facilities that pollute the air (22.1%) were located in municipalities that provide informational or financial support for ISO 14001. Facilities that emit into water had more workers and more intense emissions on average.

Tables 3.3 and 3.5 show the industry composition and each industry's average of the main variables

---

[5] 505 facilities emit into both air and water, so the total number of facilities in the two datasets is 3737.

Table 3.2: Descriptive statistics of water emitting facilities in 2001(N=663)

	mean	s. d.	max	min
Workers/100	6.88	13.5	190	0.01
Weighted sum of emissions to water/10^6 (E^W)	25.0	360	8860	6.30E-07
E^W in 2003	28.4	360	6560	0
Air emission intensity (kg*weight/persons)/10^6	0.0251	0.153	2.75	0
Water emission intensity (kg*weight/persons)/10^6	0.0809	1.25	31.8	9.84E-10
Profit/revenue (yen/thousand yen)/1000	-6.87E-04	0.0957	1.73	-0.959
Population density (persons/ha)/100	0.168	0.210	1.21	0.00140
Income per capita (thousand yen/persons)/1000	1.44	0.233	2.76	0.423
Grievances against environmental pollution/1000	0.0736	0.120	0.783	0
ISO 14001	0.585	0.493		
Support	0.198	0.398		

for water and air emitting facilities. In the water emitting facility sample, there are 11 industries, and many of these facilities also emit into the air. Equipment and supplies are the three largest. However, there are 21 industries in air emitting facility sample, and relatively fewer of these facilities also emit into water.

3.5 Estimation results

The estimation results of facilities that emit into water and air are presented separately. In the sampling routine of the SAR probit models, we utilize 200 draws for the burn-in period and then 1000 draws for the actual sampling with a 3-step Gibbs sampling of z; we set $c = (0, ..., 0)'$, $T = I_k \times 1.00e + 05$, and $d = 0.1$(the initial value) where k is the number of explanatory variables. We use a relatively faster desktop computer (with CPU Core i7 and 8 GB of memory) and R for estimation, and approximately 0.5-5 hours and 9-18 hours are required to estimate the SAR probit models with 1326 and 7182 observations, respectively.

For the WM IV and V models, we show only the models with 'm' that have the greatest likelihood(the greatest likelihood is equal to the greatest AIC and BIC). For other 'm' values, please see Table B.1-B.4 in the Appendix B.

Table 3.3: Industry composition of water emitting facilities with an average of main variables

Name of industry	Num of faci.	ISO 14001	E^W	Faci. w/ $E^A = 0$	Workers
Textile mill products Mfg.	25	40.0%	1.02	9	309.92
Pulp, paper and paper products Mfg.	27	66.7%	1.15	3	467.78
Chemical and allied products Mfg.	216	50.5%	62.6	32	294.00
Plastic products Mfg.	20	55.0%	9.17	1	430.10
Ceramic, stone and clay products Mfg.	36	58.3%	0.396	3	451.30
Iron and steel Mfg.	31	58.1%	29.9	5	1005.90
Non-ferrous metals and products Mfg.	36	33.3%	20.7	9	433.39
Fabricated metal products Mfg.	100	36.0%	6.32	54	205.73
Electrical machinery, equipment and supplies Mfg.	96	84.4%	1.63	24	1230.24
Transportation equipment Mfg.	66	69.7%	5.21	4	2410.46
Miscellaneous manufacturing industries	10	40.0%	0.268	5	254.80

Table 3.4: Descriptive statistics of air emitting facilities in 2001 (N=3579)

	mean	s. d.	max	min
Workers/100	3.63	8.11	190	0.01
Weighted sum of emissions to air/10^6 (E^A)	2.67	17.6	362	1.55E-06
E^A in 2003	1.84	12.2	396	0
Air emission intensity (kg*weight/persons)/10^6	0.0156	0.145	6.23	1.53E-09
Water emission intensity (kg*weight/persons)/10^6	0.00449	0.0848	3.53	0
Profit/revenue (yen/thousand yen)/1000	0.00382	0.0798	1.73	-2.64
Population density (persons/ha)/100	0.226	0.290	1.70	0.00113
Income per capita (thousand yen/persons)/1000	1.44	0.233	2.76	0.423
Grievances against environmental pollution/1000	0.0821	0.130	0.783	0
ISO 14001	0.414	0.493		
Support	0.221	0.415		

3.5.1 ISO 14001 adoption and performance of facilities that emit into water

Table 3.6 shows the estimation results on ISO 14001 adoption by facilities that emit into water. The number of workers in a facility has a significantly positive effect on ISO 14001 adoption at the 1% level under all specifications, whereas the water emission intensity has a negative effect at the 5% or 10% level. However, support for ISO 14001 adoption, profitability, and the characteristics of the municipality in which a facility is located do not have significant effects.

The implications of these estimation results are as follows. First, a facility with more workers is more likely to adopt ISO 14001; these findings are similar to those of Nakamura et al. (2001), Welch et al. (2002), Arimura et al. (2008) and Nishitani (2009). Second, profitability and pressure

Table 3.5: Industry composition of air emitting facilities with an average of main variables

Name of industry	Num of faci.	ISO 14001	E^A	Faci. w/ $E^W = 0$	Workers
Crude oil and natural gas production	21	47.6%	2.16	21	22.43
Food	13	46.2%	1.99	11	248.77
Textile mill products Mfg.	44	25.0%	5.05	29	219.09
Apparel and other finished products made from fabrics and similar materials Mfg.	11	0.0%	0.129	10	104.55
Lumber and wood products, except furniture Mfg.	38	78.9%	0.787	38	117.95
Furniture and fixtures Mfg.	30	60.0%	0.453	27	157.50
Pulp, paper and paper products Mfg.	93	40.9%	2.48	68	238.76
Printing and allied industries	143	17.5%	0.829	142	187.20
Chemical and allied products Mfg.	856	39.0%	5.47	657	173.39
Petroleum and Coal products Mfg.	64	54.7%	0.706	56	191.61
Plastic Products Mfg.	300	38.0%	1.15	279	173.63
Rubber Products Mfg.	108	44.4%	2.62	99	404.13
Ceramic, stone and clay products Mfg.	118	36.4%	3.80	84	294.17
Iron and steel Mfg.	86	47.7%	9.78	58	623.98
Non-ferrous metals and products Mfg.	95	36.8%	7.62	67	273.21
Fabricated metal products Mfg.	423	25.1%	0.930	377	165.61
General machinery Mfg.	239	51.0%	0.755	232	562.54
Electrical machinery, equipment and supplies Mfg.	314	67.8%	0.559	239	751.16
Transportation equipment Mfg.	395	54.2%	1.56	328	936.69
Precision instruments and machinery Mfg.	61	39.3%	1.03	51	383.30
Miscellaneous manufacturing industries	127	39.4%	0.525	121	179.11

from residents of the local community are not determinants of ISO 14001 adoption by facilities that emit into water. Third, municipalities' support for ISO 14001 adoption does not stimulate ISO 14001 adoption; facilities located in municipalities that provide support for ISO 14001 adoption are not more likely to adopt the standards than those in municipalities that do not provide support. Finally, dirty (high emission intensity) facilities are less likely to adopt ISO 14001 to improve their environmental performance, but clean facilities are more likely to adopt the standards as a sign that they are (actually) clean.

Spillovers (ν) are completely different between models. The spillover effects between large numbers of facilities are negative (WM II and III models), whereas the spillover effects between facilities with similar characteristics are positive (WM IV and V models). Except for the WM II model, where the spillover effect is significant at the 5% level, these spillover effects are significant at the 1% level. Negative spillover effects (WM II and III) might reflect that the different-sized fa-

Table 3.6: Estimation results of ISO 14001 adoption by water emitting facilities

	Dependent variable: ISO 14001 adoption				
	WM I	WM II	WM III	WM IV (m=10)	WM V (m=10)
Workers/10^2	0.142***	0.140***	0.136***	0.094***	0.109***
	(0.011)	(0.011)	(0.012)	(0.015)	(0.014)
Air emission intensity/10^6	0.025	0.03	0.043	0.075	0.029
	(0.248)	(0.246)	(0.251)	(0.271)	(0.254)
Water emission intensity/10^6	-0.200*	-0.178**	-0.224*	-0.137*	-0.147*
	(0.112)	(0.090)	(0.123)	(0.076)	(0.084)
Profit/revenue/10^3	1.04	1.00	1.10*	0.912	0.923
	(0.603)	(0.666)	(0.624)	(0.652)	(0.665)
Population density/10^3	-4.78	-4.78	-4.64	-3.81	-4.36
	(3.36)	(3.27)	(3.20)	(3.57)	(3.56)
Income per capita/10^3	0.308	0.331	0.367	0.223	0.242
	(0.359)	(0.372)	(0.360)	(0.362)	(0.376)
Support	0.007	-0.028	-0.028	-0.032	-0.016
	(0.132)	(0.140)	(0.134)	(0.140)	(0.144)
Grievances against pollution	-0.664	-0.596	-0.603	-0.612	-0.481
	(0.456)	(0.454)	(0.453)	(0.468)	(0.427)
Year dummy (2001)	-0.275***	-0.436***	-0.439***	-0.200 ***	-0.225***
	(0.083)	(0.132)	(0.122)	(0.073)	(0.078)
Constant	-13.3*	-8.53	-21.0**	-29.6 ***	-7.442
	(7.52)	(5.38)	(9.14)	(10.5)	(4.78)
ν		-0.639**	-0.669***	0.334***	0.166***
		(0.282)	(0.188)	(0.067)	(0.054)
Log-likelihood	-1474.06	-1464.54	-2245.84	-1121.92	-1099.74
BIC	-1711.33	-1705.4	-2486.7	-1362.78	-1340.60
AIC	-1540.06	-1531.54	-2312.84	-1188.92	-1166.74
LR test statistic		19.0***	-1543.6	704.3***	748.6***
Observation	1326	1326	1326	1326	1326

All of models in this table are SAR probit models. Standard errors are shown in parentheses. *, **, and *** imply that the coefficient is significantly different from zero at the 10%, 5% and 1% levels, respectively. All models are estimated with industry and prefecture dummies.

cilities have different incentives to adopt the ISO 14001 standard. Large facilities have positive attitude toward ISO 14001 adoption, while small facilities have negative attitude. The "ν"s of WM II and III are the average effects of the industry-wide and the country-wide attitude (aggregate attitude) toward ISO 14001 adoption on the individual facilities' attitudes, respectively. As mentioned previously, the ISO 14001 adoption rate of our sample is 58.5%, and the adoption rates of most industries are 40-60%. Therefore, approximately half of the facilities have attitudes toward this adoption that differ from the aggregate attitude. Negative spillover effects are likely to reflect this

Table 3.7: Estimation results of water emissions reductions

Dependent variable: Water emissions reductions (E^W_{t+1}/E^W_t)	WM I	WM II	WM III	WM IV (m=3)	WM V (m=3)
Workers	-0.367	-0.366	-0.375	-0.367	-0.367
	(1.73)	(1.73)	(1.73)	(1.73)	(1.73)
Profit/Revenue	-0.351	-0.351	-0.337	-0.352	-0.351
	(4.52)	(4.51)	(4.52)	(4.51)	(4.52)
ISO	172	172	197	176	172
	(849)	(848)	(849)	(847)	(849)
Population density	-1167	-1166	-1225	-1171	-1167
	(1746)	(1743)	(1745)	(1742)	(1746)
Income per capita	2.20	2.20	1.77	2.22	2.20
	(12.1)	(12.1)	(12.1)	(12.1)	(12.1)
Grievances against pollution	-1.90	-1.90	-2.40	-1.92	-1.90
	(10.2)	(10.2)	(10.2)	(10.1)	(10.2)
Constant	-135	-347	-185	-134	-135
	(488)	(488)	(488)	(487)	(488)
ρ		-0.990	-0.153	0.006	0.0003
		(0.000)	(0.000)	(0.000)	(0.0000)
Log-likelihood	-6647.041	-6646.352	-6646.850	-6647.038	-6647.041
BIC	-6669.780	-6672.339	-6672.837	-6673.025	-6673.028
AIC	-6654.041	-6654.352	-6654.850	-6655.038	-6655.041
LR test statistic		1.38	0.382	0.006	0.000
Observation	663	663	663	663	663

All of models in this table are fixed effects SAR models. *, **, and ***imply that the coefficient is significantly different from zero at the 10%, 5% and 1% levels, respectively.

difference in attitude toward ISO 14001 adoption.

Which effect more effectively explains the ISO 14001 adoption of water emitting facilities? To compare the specifications, we calculate BIC and AIC, according to Hepple (2004)[6]. Because the BIC and AIC difference between V and others are greater than 10, we have very strong evidence that the WM V model, or the intra-industry spillover between facilities of similar sized firms (in terms of similar revenue) has better explanatory power for ISO 14001 adoption by facilities that emit into water.

This result is quite intuitive. Because the adoption of ISO 14001 demands a large initial cost, the decision to do so is less likely to be made by the facilities than by the firms to which they belong.

---

[6] The BIC and AIC used here are BIC=log-likelihood−(the number of coefficient)×ln(Observation)/2, and AIC=log-likelihood−(the number of coefficient).

Table 3.8: Estimation results of negative binomial regressions for grievances against pollution (water emitting facilities)

	Dependent variable: Grievances against pollution			
	constant	mean	constant	mean
Percentage change in water emissions (E^W_{t+1}/E^W_t)	4.95E-6	3.10E-6		
	(3.63E-6)	(9.00E-6)		
Weighted sum of emissions into water			1.40E-10***	8.47E-11
			(5.12E-11)	(1.17E-10)
Constant	4.24***	4.24***	4.24***	4.24***
	(0.04)	(0.04)	(0.04)	(0.04)
Log-likelihood	-6650.51	-6649.96	-6650.25	-6648.34
LR test of $\alpha = 0$ or $\delta = 0$	1.4E+5***	1.4E+5***	1.4E+5***	1.4E+5***
Observation	1326	1326	1326	1326

"constant" implies that we employ a negative binomial model with constant-dispersion for estimation, and "mean" implies that we employ a negative binomial model with mean-dispersion. Standard errors are shown in parentheses. *, **, and *** imply that the coefficient is significantly different from zero at the 10%, 5% and 1% levels, respectively. all models are estimated with industry and prefecture dummies.

Similarly sized firms in the same industry (when revenue is used as a proxy for firm size) are likely to be competitors, and a model of intra-industry spillover effects between competitors can explain the data most effectively. Therefore, imitation pressure because of competition has likely contributed to the diffusion of ISO 14001 in Japan.

Table 3.7 shows the determinants of water emissions changes. We use the values of ISO 14001 predicted by ISO 14001 adoption estimation with weight matrix V. All explanatory variables are in significant at any level, and therefore, ISO 14001 adoption does not reduce water emission intensity. In addition, BIC and AIC provide little evidence that models with correlations of water emission reductions have better explanatory power than models without the correlation, although $\rho$is significant at 1% level.

Facilities with greater water emissions are likely to be the target of more grievances. Therefore, there may be an endogeneity problem in emissions reduction equations in that environmental performance affects grievances, and this endogeneity could bias the estimated coefficients. To assess whether an endogeneity problem exists, we regress the grievances variable separately on the percentage change in water emissions and on the weighted sum of water emissions. We employ a negative binomial model for the regressions. The results of these regressions are presented in Table

3.8. These results indicate that the grievances might be affected by the weighted sum of emissions but are not affected by percentage change in water emissions, which we employ as an indicator of environmental performance. Thus, there is no endogeneity due to the causal effect of environmental performance on grievances, and the estimation results of water emission reductions are not affected by such endogeneity.

3.5.2 Performance and ISO 14001 adoption of facilities that emit into air

Table 3.9 shows the estimates of ISO 14001 adoption by facilities that emit into air. The number of workers in a facility has a significantly positive effect on ISO 14001 adoption at the 1% level under all specifications, whereas grievances against pollution have a negative effect at the 10% level. However, emission intensity, support for ISO 14001 adoption, profitability, and other characteristics of the municipality in which a facility is located do not have significant effects. This result is the same as the estimation results of facilities that emit into water. Unlike facilities that emit into water, a facility located in a municipality with fewer grievances against pollution is more likely to adopt and (air) emission intensity does not. Municipalities with fewer grievances against pollution can be interpreted as cleaner ones. Facilities in such municipalities are under more pressure from a municipal government or from residents to adopt environmental practices, so the estimation result for grievances against pollution may imply that pressure from a municipal government or from residents has an impact on ISO 14001 adoption decisions.

Similarly to the facilities that emit into water, the spillover effects between large numbers of facilities are negative (WM Ⅱ and Ⅲ models), whereas the spillover effects between facilities with similar characteristics are positive (WM Ⅳ and V models). With the exception of the WM Ⅱ model, where the spillover is significant at the 10% level, these spillover effects are significant at the 1% level. Similarly to the facilities that emit into water, the negative spillover effects (WM Ⅱ and Ⅱ I) for the facilities that emit into air might reflect that the different-sized facilities have different incentives to adopt the ISO 14001 standard. This is because roughly half of the facilities have different attitude toward the adoption than the aggregate attitude, as is indicated from the fact that ISO 14001 adoption rates of our sample and of most industries are 41.4% and 40-60%, respectively.

Table 3.9: Estimation results of ISO 14001 adoption by facilities emitting into air

	Dependent variable: ISO 14001 adoption				
	WM I	WM II	WM III	WM IV (m=10)	WM V (m=10)
Workers/10^2	0.145***	0.143***	0.143***	0.097***	0.085***
	(0.005)	(0.006)	(0.005)	(0.016)	(0.006)
Air emission intensity/10^6	0.148	0.169	0.159	0.163	0.165
	(0.129)	(0.127)	(0.125)	(0.130)	(0.125)
Water emission intensity/10^6	-0.121	-0.121	-0.101	-0.139	-0.118
	(0.109)	(0.196)	(0.103)	(0.107)	(0.102)
Profit/revenue/10^3	0.159	0.079	0.064	-0.05	0.061
	(0.227)	(0.230)	(0.232)	(0.220)	(0.233)
Population density/10^3	-1.72	-1.69	-1.74	-1.95	-1.20
	(1.15)	(1.13)	(1.07)	(1.24)	(1.17)
Income per capita/10^3	-0.11	-0.112	-0.121	-0.216	-0.129
	(0.112)	(0.115)	(0.112)	(0.123)	(0.121)
Support	0.069	0.071	0.074	-0.054	0.059
	(0.049)	(0.049)	(0.047)	(0.046)	(0.048)
Grievances against pollution	-0.293	-0.305*	-0.293*	-0.235	-0.355**
	(0.169)	(0.166)	(0.161)	(0.182)	(0.171)
Year dummy (2001)	-0.209***	-0.316***	-0.375***	-0.130 ***	-0.131***
	(0.319)	(0.072)	(0.050)	(0.035)	(0.028)
Constant	-15.6***	5.12**	-11.9**	-54.5 **	-3.02**
	(4.67)	(2.17)	(5.27)	(21.4)	(1.26)
ν		-0.544*	-0.847***	0.430***	0.417***
		(0.305)	(0.097)	(0.090)	(0.025)
Log-likelihood	-6018.23	-6016.8	-6027.13	-6352.29	-6337.96
BIC	-6359.96	-6358.53	-6368.86	-6694.01	-6679.68
AIC	-6095.23	-6093.80	-6104.13	-6429.29	-6414.96
LR test statistic		2.86*	-17.8	-668.1	-638.5
Observation	7158	7158	7158	7158	7158

All of models in this table are SAR probit models. Standard errors are shown in parentheses. *, **, and *** imply that the coefficient is significantly different from zero at the 10%, 5% and 1% levels, respectively. All specifications are estimated with industry and prefecture dummies.

Among the models with a spatial correlation, WM II has the greatest BIC and AIC. Compared with the WM I model, the WM II model has an AIC that is greater by approximately 1 but a BIC that is smaller by approximately 3. Therefore, we do not have positive evidence that the specification with the spillover effect performs better.

The determinants of air emission changes are presented in Table 3.10. We use values of ISO 14001 predicted by ISO 14001 adoption estimation with Weight Matrix I (no spatial correlation model). All of the explanatory variables are insignificant at all levels, and therefore, ISO 14001 adoption

Table 3.10: Estimation results of air emissions reductions (SAR)

	Dependent variable: Air emissions reductions (E^A_{t+1}/E^A_t)				
	WM I	WM II	WM III	WM IV (m=5)	WM V (m=10)
Workers	0.0248	0.0248	0.0276	0.0249	0.0256
	(0.0528)	(0.0528)	(0.0528)	(0.0515)	(0.0526)
Profit/Revenue	0.0163	0.0163	0.068	0.0069	0.0189
	(0.100)	(0.100)	(0.100)	(0.098)	(0.100)
ISO	-13.1	-13.1	-14.7	-13.8	-13.3
	(20.6)	(20.6)	(20.6)	(20.1)	(20.5)
Population density	6.41	6.41	5.81	5.21	5.97
	(26.7)	(26.7)	(26.7)	(26.1)	(26.6)
Income per capita	-0.178	-0.178	-0.183	-0.169	-0.176
	(0.233)	(0.233)	(0.232)	(0.227)	(0.232)
Grievances against pollution	0.144	0.144	0.13	0.131	0.142
	(0.202)	(0.202)	(0.201)	(0.197)	(0.201)
Constant	6.72	16.4	3.69	6.06	6.36
	(10.8)	(10.8)	(10.7)	(10.5)	(10.7)
ρ		-0.990***	0.318***	0.110***	0.0390***
		(0.000)	(0.000)	(0.000)	(0.000)
Log-likelihood	-25426.65	-25426.96	-25423.02	-25419.32	-25426.18
BIC	-25455.29	-25459.69	-25455.75	-25452.06	-25458.91
AIC	-25433.65	-25434.96	-25431.02	-25427.32	-25434.18
LR test statistic		-0.620	7.26***	14.7***	0.940
Observation	3579	3579	3579	3579	3579

All of models in this table are fixed effects SAR models. *, **, and *** imply that the coefficient is significantly different from zero at the 10%, 5% and 1% levels, respectively.

does not reduce air emission intensity as it does water emission intensity. However, ρ is significant at the 1% level for all models with spillover effects, and BIC and AIC provide positive evidence that the Weight Matrix IV model has better explanatory power than one without the correlation.

This result implies that the percentage change in air emissions of similarly sized facilities in the same industry are correlated; i.e., a facility reduces its air emissions if similarly sized facilities do so. This correlation may reflect technology diffusion; similarly sized facilities in the same industry are likely to have similar technology. In contrast with air emitting facilities, the low explanatory power of spillover models among similarly sized firms and facilities for water emission reductions may reflect the fact that water emitting facilities have different technology even though they are similar in size.

Is ρ for the air emission reduction equation (particularly, in WM IV model) significant because

Table 3.11: Estimation results of air emissions reductions (SAR and SAC)

	Dependent variable: Air emissions reductions (E^A_{t+1}/E^A_t)		
	SAC FS WM I m=5	SAR FS WM II m=5	SAC FS WM II m=5
Workers	0.0249	0.0248	0.0248
	(0.0515)	(0.0515)	(0.0515)
Profit/Revenue	-0.00693	-0.00686	-0.00679
	(0.0980)	(0.0980)	(0.0980)
ISO	-13.8	-13.0	-13.0
	(20.1)	(20.0)	(20.0)
Population density	5.21	5.15	5.15
	(26.1)	(26.1)	(26.1)
Income per capita	-0.169	-0.169	-0.169
	(0.227)	(0.227)	(0.227)
Grievances against pollution	0.131	0.131	0.131
	(0.197)	(0.197)	(0.197)
Constant	6.06E	5.99	6.00
	(8.69)	(10.5)	(8.71)
ρ	0.110^{***}	0.109^{***}	0.109^{***}
	(0.027)	(0.000)	(0.027)
τ	-0.999		-0.999
	(1.41)		(1.41)
Log-likelihood	-25418.62	-25419.34	-25418.67
Observation	3579	3579	3579

*, **, and *** imply that the coefficient is significantly different from zero at the 10%, 5% and 1% levels, respectively. FS WM I and II imply we use values of ISO 14001 predicted by the first stage estimation with weight matrix I and II, respectively.

emission reductions are actually correlated or only because the errors are correlated? To examine this question, we employ Spatial Auto-Correlation(SAC) models to examine whether ρ (correlation in air emissions reductions between facilities) for the air emission reduction equation is significant because emission reductions are actually correlated or only because errors are correlated. A typical SAC model is as follows:

$$\begin{aligned} y &= \rho W_1 y + X\beta + u \\ u &= \tau W_2 u + \varepsilon \quad \varepsilon \sim N(0, \sigma I_n) \end{aligned} \tag{3.11}$$

W_1 and W_2 can be the same or different. If $\rho = 0$, this model is the same as a Spatial Error Model (SEM). First, we estimate SEMs to select W_2 from the five weight matrices used for the SAR

Table 3.12: Estimation results of negative binomial regressions for grievances against pollution (air emitting facilities)

	Dependent variable: Grievances against pollution			
	constant	mean	constant	mean
Percentage change in water emissions (E^A_{t+1}/E^A_t)	-3.59E-6	-1.62E-5		
	(4.67E-5)	(6.24E-5)		
Weighted sum of emissions into water			1.81E-9***	1.80E-12
			(6.67E-10)	(1.18E-09)
Constant	4.36***	4.36***	4.35***	4.36***
	(0.02)	(0.02)	(0.02)	(0.02)
Log-likelihood	-36486.11	-36486.06	-36483.01	-36486.11
LR test of $\alpha = 0$ or $\delta = 0$	8.8E+5***	8.8E+5***	8.8E+5***	8.8E+5***
Observation	7158	7158	7158	7158

"constant" implies that we employ a negative binomial model with constant-dispersion for estimation, and "mean" implies that we employ a negative binomial model with mean-dispersion. Standard errors are shown in parentheses. *, **, and ***imply that the coefficient is significantly different from zero at the 10%, 5% and 1% levels, respectively. all models are estimated with industry and prefecture dummies.

models. Because the likelihood of the WM I model is the greatest, we select WM I for W_2. Then, we estimate the SAC models by setting W_2 to weight matrix II. In addition, we estimate equations for the percentage change in air emissions with values of the ISO 14001 adoption predicted by the model with WM II.

Table 3.11 shows estimates of the SAC model with the value of ISO 14001 predicted by ISO 14001 adoption estimation with weight matrix I in the first column, and the SAR and SAC model with the value of ISO 14001 predicted by ISO 14001 adoption estimation with weight matrix II in the second and third columns, respectively. We use weight matrix IV for all these models. According to Table 3.11 and the last column of Table 3.10, ρ is significant and its value changes little, even with error correlation. This result indicates that percentage changes in emission reductions are actually correlated.

As with the water-emitting facilities, we assess whether an endogeneity problem arises for the air-emitting facilities because of the causal effect of environmental performance on grievances by regressing the grievances variable separately on the percentage change in air emissions and on the weighted sum of air emissions. Table 3.12 presents the results of these regressions. These results indicate that the estimation results of air emissions reductions are not affected by such endogeneity.

3.6 Conclusion

In the OECD surveys, many facilities evaluate the practices “other facilities like ours is adopting ” as “very important” or “important”as a motivation for adopting an EMS and environmental practices. This chapter examines such externalities in ISO 14001 adoption and environmental performance, and it also examines the determinants of ISO 14001 adoption and environmental performance by employing spatial econometric estimation methods. We focus on Japanese facilities because Japan has ranks first worldwide with the most facilities that have adopted ISO 14001.

We find the following results for ISO 14001 adoption. First, the number of workers in a facility has significantly positive effects on ISO 14001 adoption, as was found in previous studies on ISO 14001. Second, facilities that emit into water with lower water emission intensity are more likely to adopt ISO 14001. Third, the ISO 14001 adoption by facilities that emit into water can be more effectively explained by the intra-industry spillover effects between facilities of firms with similar revenue than other spillover effects, whereas the intra-industry spillover effects cannot more effectively explain the adoption of facilities that emit into air. Because we employ revenue as a proxy of firm size, similar-sized firms make similar decisions on ISO 14001 adoption. In addition, this result may indicate that the ISO 14001 adoption decisions of competitors influence each other as similar-sized firms in the same industry are likely to be competitors,

The results of this chapter suggest that ISO 14001 adoption does not have a significant impact on environmental performance or on other factors. However, the percentage change in air emissions, which was used as a proxy for environmental performance, is correlated between similar sized facilities in the same industry. This correlation might reflect technology diffusion because similar sized facilities in the same industry are likely to use similar technology. Our estimation results suggest the possibility that similar-sized firms or facilities implement similar voluntary environmental actions and that there are unlikely spillover effects of voluntary environmental actions between facilities and workers with various characteristics. Our finding son spillover effects suggest that governments might be able to pursue voluntary programs of industry-specific or narrower targets in cooperation with industry associations.

We focus on intra-industry spillover effects of voluntary environmental actions in Japan. However, geographical spillover effects of voluntary environmental actions might also exist and may be correlated with other spillover effects. Therefore, it would be interesting to examine these two types of spillover effects simultaneously, although an estimation including two spillover effects would be much more complicated and computation intensive. Examining multiple simultaneous spillover effects remains a topic for future research.

Chapter 4

Taxes versus Quotas in Lobbying by a Polluting Industry with Private Information on Abatement Costs

4.1 Introduction

Two of the main causes of inefficiency of environmental regulations are informational issues about polluters' abatement cost and political pressure by polluting industries when the regulator can specify how much particular industries are permitted to pollute. Many researchers have studied these two causes and their impacts on environmental regulation.

For informational issues, Weitzman (1974) first compared taxes and quotas in the presence of uncertainty about marginal costs and benefits in a general context. Since Weitzman, many researchers have compared these two instruments for environmental protection under uncertainty about abatement cost in different settings[1]. Moledina et al. (2003) and Costello and Karp (2005) focus on asymmetric information about abatement costs and compare taxes and permits in a dynamic setting. While the firms take advantage of asymmetric information and behave strategically in the model of Moledina et al. (2003) the firm in Costello and Karp (2005). is non-strategic. Other papers have modeled a pollution control problem under asymmetric information as a principal-agent or mechanism design problem. Dasgupta et al. (1980) applied a Groves mechanism to a multiple-

---

[1] For example, several papers compare two instruments for stock pollutant control (Hoel and Karp (2001), Hoel and Karp (2002), Karp and Zhang (2002), and Newell and Pizer (2003)). Quirion (2004) analyzes a case with pre-existing distortionary taxes, and Kaplow and Shavell (2002)analyze a case in which a regulator can impose a nonlinear tax

polluter regulation problem, while Lewis (1996) and other studies employed the principal-agent framework. In these frameworks, a government implements a combination of pollution quotas and taxes (or subsidies).

Research on the political economy of environmental policies began recently relative to research on informational issues. By applying to environmental policies a lobbying model developed byGrossman and Helpman (1994), Fredriksson (1997) and Aidt (1998) independently examine how pollution taxes are affected by lobbying groups. Damania (2001) shows if firms can political influence the goverrnment, then they underinvest in clean technologies to make lobby against environmental policy effective. Conconi (2003) analyzes how a green lobby affects trade and environmental policies in two large economies with transboundary pollution. Barbier et al. (2005) theoretically and empirically examine how dynamic resource conservation policies are affected by industry lobbies. Yu (2005) analyzes a model where lobbying groups affect public environmental awareness by advertising before they lobby the government. These papers assume that a government knows the pollution abatement cost of industry.

The above studies take into account only one of the two causes of inefficient regulation. However, governments usually face both informational and political economy issues. This chapter compares taxes and quotas when a government faces both issues or is affected by an industry lobby with private information about its abatement cost. Especially, we examine the effects of private information about abatement cost on lobbying activity, environmental regulation and social welfare and show how different these effects are under tax and quota systems. To do so, we incorporate the principal-agent relationship with an informed principal and common-value private information examined by Maskin and Tirole (1992)[2] into a lobby model of the Grossman and Helpman type.

Our results differ from the literature. First, our result on social welfare is completely different from the results in the literature on taxes versus quotas under imperfect information. In the literature on "tax versus quota", a tax is basically considered better than a quota if the marginal abatement cost

---

[2] In general, we can classify the principal's private information into one of two forms based on whether the principal's private information is an argument of the agent's objective function. If it is an argument, the private information is classified as common value. Otherwise, the private information is classified as private value. Because a government of this chapter's model, an agent, takes into account abatement cost, like most models of lobbying on environmental policy, private information is an argument of the agent's objective function. For this reason, we employ a common-value model.

(MAC) curve's slope is steeper than the marginal damage (MD) curve's slope. In our model, however, quotas are socially preferred (1) when the government weighs political contributions much more strongly than it does social welfare or (2) when the MAC curve's slope is steeper than the MD curve's slope. The industry generally has a stronger incentive to, and actually does, relax regulations under taxes than under quotas due to tax payments. This finding is particularly the case when the government is easily affected by political contributions because paying political contributions for low tax rates is cheaper than paying tax under such a case.

The industry's benefit from lobbying is larger under taxes than it is under quotas if the slope of the MAC curve is steep (relative to the MD curve). Consequently, quotas are better than taxes when the MAC curve's slope is steep. Because a steep slope implies that a price changes more than the quantity does, a reduction in the abatement target lowers the tax rate and tax payments more than the (actual) abatement does when the MAC curve's slope is steep. Therefore, in such a case, the industry's cost savings due to the reduction of the abatement target by one unit is greater under taxes than it is under quotas. Therefore, the government is under high political pressure from the industry, and it sets a low tax rate.

The impact of private information on social welfare under our model might also be different from that under standard regulation models. It is well known that if a polluting industry cannot politically influence environmental regulations (standard regulation models), a government has to pay an information rent to some type(s) of industry that has an incentive to pretend that it is of another type(s). However, if the industry can do so, it may have to reveal its own type (on abatement cost), with some cost. For example, when the tax is implemented, a low-cost industry (i.e., low abatement cost) will face a lower tax rate and lower political contribution than a high-cost industry (because the low-cost industry will abate enough even if its tax rate is low relative to that of the high-cost industry). Therefore, the high-cost industry has an incentive to pretend to be the low-cost industry, while the low cost industry evades higher tax rate and political contribution by revealing its type. Relative to the high-cost industry, the low-cost industry can endure a high tax rate because it does not emit much, but the high-cost industry has higher willingness to pay for a low tax rate compared to the low-cost industry. The low-cost industry has to accept a higher tax rate in order to reveal its type. The higher tax rate constitutes a cost to the industry, but it is a benefit for society (provided that the tax rate is not too high). Thus, private information in a case with political influence of the polluting industry has different impacts compared to private information in a case without political influence.

Actually, what happens under a tax depends on whether the slope of MAC curve is steeper than the respective slopes of the (actual) MD curve and the weighted MD curve (MD curve multiplied by government's weight on social welfare)[3]. We show that if the MAC curve's slope is steeper than the weighted MD curve's slope, the private information might improve social welfare by the mechanism discussed in the last paragraph. However, if the MAC curve's slope is steeper than the MD curve's slope, the low-cost industry might not be able to reveal its type. Because the difference in tax rates between the low- and high-cost industries is great in such a case (as price changes more than quantity does), the high-cost industry has a strong incentive to pretend to be the low-cost industry in order to save abatement cost and reduce its tax payment. Therefore, it is very hard or impossible for the low-cost industry to reveal its type, and a separating equilibrium where the government can differentiate the industry's type might not exist.

In contrast to the case of taxes, private information does not improve social welfare when quotas are implemented. This is because that the high-cost industry might have to make the government set a higher quota than it would under complete information to reveal its cost information. However, a separating equilibrium always exists, unlike in the case of taxes. The quota for the high-cost industry under a complete information case is higher than that for the low-cost industry, and the high-cost industry is more in favor of a high quota compared to the low-cost industry. Therefore, by making the government set a higher quota than would be set under the complete information case, the high-cost industry can reveal its type to the government. Thus, private information reduces the comparative disadvantage of taxes compared to quotas when the government cares little about social welfare.

Although environmental groups also lobby the government, it appears that analyzing only the industry lobby with private information is not unreasonable. During the 1990s, contributions to the U.S. Congress from electric utilities and the oil and gas industry alone were roughly 10 times larger

---

[3] "The MAC curve slope is steeper than the MD curve slope or the weighted MD curve slope" means that the abatement cost changes more than the damage or weighted damage caused by a one-unit change in emissions.

than the contributions from environmental organizations according to data from the Center for Responsive Politics. In addition, there were no political contributions from pro-environmental policy groups in 10 states in the US, and the total contribution from the energy industry was at least 10 times larger than that from pro-environmental policy groups in most states during 2003-2006 [4].

In this chapter, the industry lobby group indirectly transmits its private information to the government. Therefore, the paper might fall within an informational lobbying literature that analyzes the transmission of information from lobby group(s) to a policy-maker. Some papers in the literature analyze cases with two (or three) policy alternatives (e.g., Austen-Smith and Wright (1994), Lohmann (1995), Bennedsen and Feldmann (2006), Dahm and Porteiro (2008)). Other papers examine the case when the government chooses a policy from a continuous set of alternatives, as in this chapter (e.g., Austen-Smith (1995), Austen-Smith and Banks (2002)). A continuous set of alternatives (regulation levels) is preferable when analyzing how the regulation level is affected by lobbying. However, in papers with continuous sets of alternatives, the policy must be very abstract to enable the analysis of strategic transmission of information. In addition, according to the form of the utility function, every type has a different preference for policy in these papers, and every type of industry wants a lower emission tax or a higher emission quota. Thus, the settings of these papers are not suitable to analyze lobbying for environmental policies by polluting industries possessing private information.

Finally, we have to mention Boyer and Laffont (1999) who, like us, examine environmental policies under a case with political economy and informational issues. They compare the incentive mechanism (abatement level and transfer) with a single abatement level based on expected abatement cost when a monopoly has private information on pollution abatement cost and political majorities representing different stakes chooses the instrument. They show that the incentive mechanism may not be desirable if there are more stakeholders in the monopoly in a society or if informational rents to the stakeholders are large. However, they focus only on quantity regulations (direct quantity control regulations), whereas this chapter compares quantity and price instruments (direct and indirect control regulations).

---

[4] There are two exceptions. In Alabama, contributions from the energy industry amounted to \$811,300, whereas contributions from pro-environmental policy groups amounted to \$421,409. In Oregon, \$993,038 was contributed from the energy industry and \$260,278 from pro-environmental policy groups. See Moore (2007) for further detail.

This chapter is organized as follows. Section 2 describes the model environment. Section 3 argues the complete information case, while section 4 argues the incomplete information case. Section 5 compares the consequences of a tax with those of a quota using numerical methods. We conclude in section 6.

4.2 Setup

There are two types of industries, one with low abatement cost and one with high abatement cost, denoted by $i = L, H$. Let the abatement cost of type i be

$$C_i(e_i) = C(e_i, \bar{e}_i) = \begin{cases} \frac{1}{2}c(\bar{e}_i - e_i)^2 & if\ \bar{e}_i \geq e_i \\ 0 & if\ \bar{e}_i < e_i \end{cases} \tag{4.1}$$

for $i = L, H$ where e_i is emission by type i and $\bar{e}_i$ can be private information held by the industry. c, $\bar{e}_i$, and $\bar{e}_i - e_i$ can be interpreted as the slope of the marginal abatement cost (MAC) curve, natural emission of type i and pollution abatement of type i, respectively.

The industry of type i makes a political contribution, Ω_i, to make an environmental regulation (quota or tax) more lenient. The total cost of the industry of type i is $C_i(e_i(q)) + \Omega_i(q)$ under a quota and is $C_i(e_i(t)) + te_i(t) + \Omega_i(t)$ under a tax, where q is the quota and t is the tax rate.

Damage by pollution is denoted by $D(e) = \frac{1}{2}e^2$. Therefore, a benevolent policymaker set an environmental regulation to minimize the social cost, total pollution abatement cost and damage by pollution, $SC_i(e_i) = D(e_i) + C_i(e_i)$. However, policy-making will be influenced by lobbying by the industry. We follow Grossman and Helpman and assume that the objective function of the government is $\Omega_i(t) - \alpha SC_i(e_i(t))$ when it implements a tax and $\Omega_i(q) - \alpha SC_i(e_i(q))$ when implementing a quota, where $\alpha > 0$ is the exogenously given weight of social welfare relative to the contribution.

Political process

There are two stages in the political process between the government and industry. In the first stage, the industry offers the government a contingent contribution schedule $\Omega_i(t)$ in the tax case and $\Omega_i(q)$ in the quota case. $\Omega_i(t)$ and $\Omega_i(q)$ are continuously differentiable functions mapping from T to R_+ and from Q to $R_+$where $T = [0, c\bar{e}_H]$ and $Q = [0, \bar{e}_H]$ are one-dimensional tax and quota choice set, respectively. In the second stage, the government selects a regulation level (tax rate or quota) and receives from the industry the contribution associated with the selected regulation level.

4.3 Complete information case

We characterize the equilibria under the complete information cases as follows.

If the government implements a tax, $(\{\Omega_i^C, t_i^C\}_{i=L,H})$ is the Subgame Perfect Nash Equilibrium (SPNE) if

(i) t_i^C maximizes $\Omega_i^C(t_i) - \alpha SC_i(t_i)$

(ii) t_i^C minimizes $C_i(e_i(t_i)) + t_i^C e_i(t_i) + \Omega_i^C(t_i)$ s.t. $\Omega_i^C(t_i^C) \geq \alpha[SC_i(t_i^C) - SC_i(t_i^*)]$ where t_i^* minimizes $SC_i(t)$ for all i.

If the government implements quota, $(\{\Omega_i^C, q_i^C\}_{i=L,H})$ is the SPNE if

(i) q_i^C maximizes $\Omega_i^C(q_i) - \alpha SC_i(q_i)$

(ii) q_i^C minimizes $C_i(e_i(q_i^C)) + \Omega_i^C(q_i)$ s.t. $\Omega_i^C(q_i^C) \geq \alpha[SC_i(q_i^C) - SC_i(q_i^*)]$ where q_i^* minimizes $SC_i(q_i)$ for all i.

Condition (i) says that the government sets the regulation to maximize its objective function given contribution schedule of the industry. Condition (ii) implies that the equilibrium regulation minimizes the industry's cost subject to the participation constraint of the government. Contribution schedules that minimize the industry's cost must satisfy the participation constraint with equality. This means that at equilibrium, the government gets the same payoff as it gets when it does not reject the industry's offer. In other words, the industry has the first mover's advantage.

Both conditions characterize the contribution schedule. It is required that the equilibrium contribution schedule be $\Omega_i(t_i) \leq \alpha SC_i(t_i) - \alpha SC_i(t_i^*)$ for all t_i on T when taxes are implemented. $\Omega_i(q_i) \leq \alpha SC_i(q_i) - \alpha SC_i(q_i^*)$ for all q_i on Q when a quota is implemented.

4.3.1 Equilibrium tax rate

t_i^Cminimizes $C_i(e_i(t_i)) + t_i e_i(t_i) + \alpha[SC_i(e_i(t_i)) - SC_i(t_i^*)]$ on T. Because the best response to t_i is $e_i(t_i) = \bar{e}_i - t_i/c$ (if $t_i \leq c\bar{e}_i$)[5],

$$\begin{aligned} C_i(e_i(t_i)) + t_i e_i(t_i) + [SC_i(e_i(t_i)) - SC_i(t_i^*)] &= \frac{1}{2}\left\{2t(\bar{e}_i - \frac{t}{c}) + \frac{t}{c}^{2} + \alpha[\frac{t^2}{c} + (\bar{e}_i - \frac{t}{c})^2 - \frac{c}{1+c}\bar{e}_i^2]\right\} \\ &= \frac{1}{2}\left[\frac{(\alpha-1)c+\alpha}{c^2}t^2 + 2\bar{e}_i\left(1 - \frac{\alpha}{c}\right)t + \frac{\alpha}{1+c}\bar{e}_i^2\right]. \end{aligned} \tag{4.2}$$

If $(\alpha - 1)c + \alpha < 0$, $t_i = 0$ is optimal[6]. If $(\alpha - 1)c + \alpha \geq 0$, then

$$t_i^C = \begin{cases} \frac{\bar{e}_i(\alpha - c)c}{(\alpha-1)c+\alpha} & if\ \alpha > c \\ 0 & otherwise. \end{cases} \tag{4.3}$$

When $(\alpha - 1)c + \alpha < 0$, $\alpha < c$ always holds. Hence, (4.3) characterizes the equilibrium tax rate. The slope of marginal damage (MD) is normalized to 1, but if the slope of the MD is d, then $t_i^C = \frac{\bar{e}_i(\alpha d - c)c}{(\alpha-1)c+\alpha d}\ if\ \alpha d > c$ or $t_i^C = 0$ otherwise. Therefore, we can interpret $\alpha < c$ as a case in that the slope of the MAC curve is steeper than that of a weighted MD curve which is equal to the MD curve multiplied by the government's weight given to social welfare.

High α can be interpereted as high environmental awareness or low political pressure from the polluting industry because a lobbying group is a polluting industry only. Hence, we can interpret the above results as indicating that the introduction of a pollution tax is blocked or postponed by political pressure from the polluting industry when envrionmental awareness is low or political

---

[5] From F.O.C, $c(\bar{e}_i - e_i) = t$. And, $e_i(t_i) = 0$ if $t_i > c\bar{e}_i$.

[6] In this case, $t_i = 0$ or $t_i = c\bar{e}_i$ might be the solution to the above problem. Because $C_i(e_i(0)) + \alpha SC_i(e_i(0)) = \alpha\bar{e}_i^2, C_i(e_i(\bar{e}_i)) + c\bar{e}_i e_i(c\bar{e}_i) + \alpha\ SC_i(e_i(c\bar{e}_i)) = (\alpha+1)c\bar{e}_i^2$, and $\alpha < (1-\alpha)c < c$, $t_i = 0$ is the solution.

pressure from the industry is high.

4.3.2 Equilibrium quota

q_i^Cminimizes $C_i(e_i(q_i)) + \alpha[SC_i(e_i(q_i)) - SC_i(q_i^*)]$ on Q. Because the best response to q_i is $e_i(q_i) = q_i$ (if $q_i \leq \bar{e}_i$)[7],

$$\begin{aligned} C_i(e_i(q_i)) + \alpha[SC_i(e_i(q_i)) - SC_i(q_i^*)] &= \frac{1}{2}[c(\bar{e}_i - q_i)^2 + \alpha\{q_i^2 + c(\bar{e}_i - q_i)^2 - \frac{c}{1+c}\bar{e}_i^2\}] \\ &= \frac{1}{2}[((\alpha+1)c + \alpha)q_i^2 - 2(\alpha+1)c\bar{e}_i q_i + \frac{c(1+c+\alpha c)}{1+c}\bar{e}_i^2]. \end{aligned} \tag{4.4}$$

Therefore, the optimal quota is

$$q_i^C = \frac{(\alpha+1)c}{(\alpha+1)c+\alpha}\bar{e}_i. \tag{4.5}$$

In order to compare social welfare, we compare the respective emission levels under both taxes and quotas. Because emission in both cases is higher than is socially optimal, a lower emission level implies higher welfare. If government is benevolent (i.e.,α is infinity), then taxes and quotaes are equivalent. But, if $\alpha < c$ or if

$$\begin{aligned} e_i(t_i^C) \geq q_i \Leftrightarrow \frac{\alpha c}{(\alpha-1)c+\alpha}\bar{e}_i &\geq \frac{c(1+\alpha)}{c+\alpha c+\alpha}\bar{e}_i \\ \Leftrightarrow c &\geq \frac{\alpha}{1+\alpha}, \end{aligned} \tag{4.6}$$

then the welfare under a quota is higher than the welfare under a tax. From (4.6), it is likely that a quota will be better than a tax in terms of social welfare when the government places much more weight on contributions than it does on social welfare. This result is quite intuitive because a tax has a negative income distribution effect on the industry and tax payments as well as on marginal abatement cost. Another implication of (4.6) is that a quota is always better than a tax when the slope of marginal abatement cost (MAC) is steeper than that of marginal damage (MD) ($c \geq 1$). To

---

[7] $e_i(q_i) = \bar{e}_i$ if $q_i > \bar{e}_i$.

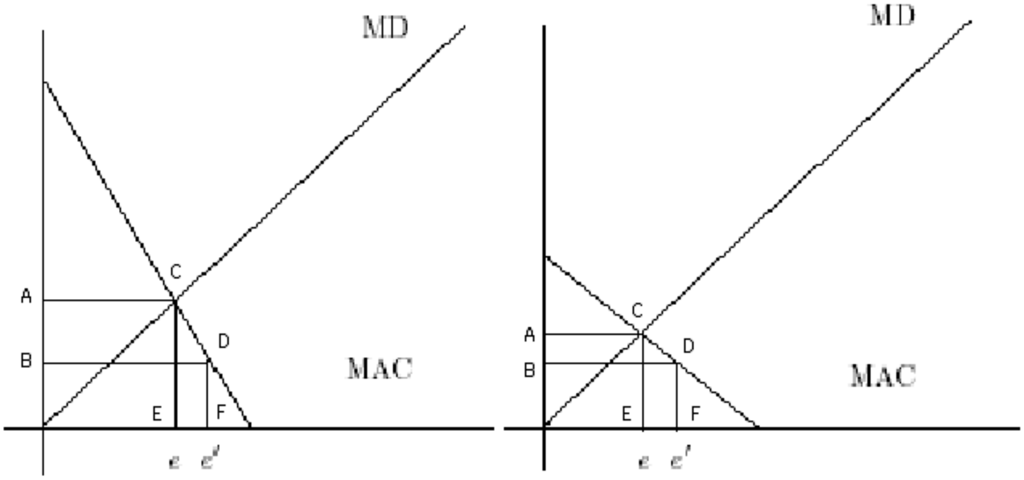


Figure 4.1: Cost savings under the tax and the quota when the slope of MAC is steeper (left) and flatter (right) than the slope of MD

understand this, consider the cost savings under a tax and under a quota when an emission target increases by one unit. Figure 1 shows the cost savings under the tax and under the quota when the slope of MAC is steeper (left) and flatter (right) than that of MD. The cost savings under the tax is area ABDC in both the left and right graphs, while cost savings under the quota is area CDFE. The cost savings under the tax is likely greater than that under the quota if the slope of MAC is steeper than that of MD. From the above effects, we obtain (4.6)

Because a tax has a negative income distribution effect on the industry, the industry prefers quotas over taxes.

Proposition 4.1. *The industry always prefers quotas over taxes under complete information cases.*

Proof: See appendix C.4.

4.4 Incomplete information case

In the incomplete information case, we cannot employ SPNE as an equilibrium concept; therefore, we employ Perfect Bayesian Equilibrium (PBE) like standard incomplete information games. As

Chapter13 of Mas-colell et al. (1995), a set of strategies and a belief function $\mu(\Omega) \in [0,1]$ (or $\mu(\Omega) \in [0,1]$) representing the government's assessment of probability that the industry is of low abatment cost after observing the contribution schedule Ω is a PBE if

(i) The industry's strategy is optimal given the government's strategies.

(ii) The belief function $\mu(\Omega)$ is derived from the industry's strategy using Bayes' rule where possible.

4.4.1 Tax

Equation (4.3) implies that the tax rate of the low-cost industry is weakly lower than that of the high-cost industry. The political contribution under some tax rate is proportional to the difference in social cost between that tax rate and the socially optimal tax rate. Relative to the high-cost industry, this difference is small if the industry is of low cost. The political contribution of the low-cost industry for some tax rate (lower than the socially optimal rate) is smaller than that of the high-cost industry. Therefore, the high-cost industry has an incentive to pretend to have low costs. In the next sub-subsection, we will present the case in which the high-cost industry cannot pretend to have low costs (separating equilibrium). Then, we will discuss the case in which the government cannot tell the industry's type (pooling equilibrium).

4.4.1.1 Separating equilibria

Under a separating equilibrium, a government differentiates the industry's type. In other words, the high-cost industry has no incentive to pretend to be the low-cost industry and incentive compatibility (IC) conditions hold under a separating equilibrium. We can write IC conditions as

$$C_H(t_H^C) + t_H^C e_H(t_H^C) + \Omega_H(t_H^C) \leq C_H(t_L) + t_L e_H(t_L) + \Omega_L(t_L) \tag{4.7}$$

$$C_L(t_L) + t_L e_L(t_L) + \Omega_L(t_L) \leq C_L(t_H^C) + t_H e_L(t_H^C) + \Omega_H(t_H^C), \tag{4.8}$$

different types have to accept to reveal their typeand PC conditions (conditions under which the government accepts the offer and chooses t_i for $i = L, H)$ as $\Omega_i(t_i) \geq \alpha[SC_i(t_i) - SC_i(t_i^*)]$ and $\Omega_i(t) \leq \Omega_i(t_i) + \alpha[SC_i(t) - SC_i(t_i)]$ for all $t = \in T$ and $i = L, H$.

Let t_L^{HIC} be a tax rate such that (4.7) holds with equality and $\Omega_L(t_L^{HIC}) = \alpha[SC_L(t_L^{HIC}) - SC_L(t_L^*)]$ and let be t_L^{LIC} a tax rate such that (4.8) holds with equality and $\Omega_L(t_L^{HIC}) = \alpha[SC_L(t_L^{HIC}) - SC_L(t_L^*)]$. t_L^{HIC} is the lowest tax rate at which the high-cost industry has no incentive to pretend that it is of low cost, whereas t_L^{LIC} is the highest tax rate at which the low-cost industry has an incentive to reveal its type[8]. Therefore, if $t_L^{LIC} > t_L^{HIC}$, then a separating equilibrium exists. The following proposition states the necessary and sufficient condition for the existence of a separating equilibrium.

Proposition 4.2. $t_H^C \leq \hat{t}$ *if and only if a separating equilibrium exists where* $\hat{t}$ *is a tax rate such that* $SC_L(\hat{t}) - SC_L(t_L^*) = SC_H(\hat{t}) - SC_H(t_H^*)$.

Proof: See appendix C.5.

Figures 4.2 and 4.3 give a graphical proof of the existence of a separating equilibrium. Figure 4.2 illustrates a case where a separating equilibrium exists, and figure 4.3 represents a case with no separating equilibrium. The PC_L and PC_H curve are a set of (t, Ω) that satisfy PC conditions for low-cost and high-cost industries with equality in both figures 4.2 and 4.3, respectively. Because the government prefers high contributions, the areas above PC_L and PC_H satisfies PC conditions for the low-cost and high-cost industries, respectively. The IC_L and IC_H curves are isocost curves for the low- and high- cost industries (the same cost as in the case when the tax rate is t_H^C and the contribution is Ω_H^C), respectively. Becuse the industry likes low tax rates and small contributions, the low-cost industry prefers combinations of taxes and contributions in the area below IC_L over (t_H^C, Ω_H^C) and the high-cost industry prefers points below IC_H over (t_H^C, Ω_H^C). Relative to the high cost industry, the low cost industry can endure a high tax rate because it emits much less. Therefore, IC_L intersects the x-axis at a higher tax rate than IC_H does. Under separating equilibria, the low

---

[8] The low cost industry can reduce its cost by decreasing Ω_L and increasing t_L with IC for the high-cost industry, holding with equality if $\Omega_L(t_L) > \alpha[SC_L(t_L) - SC_L(t_L^*)]$ This is because the increase in the sum of the abatement cost and the tax payment of the low-cost industry is smaller than that of the high-cost industry. Of course, whether $\Omega_L(t_L) = \alpha[SC_L(t_L) - SC_L(t_L^*)]$ is an equilibrium contribution depends on the government's belief about the industry's cost.

cost industry prefers a combination of tax rate and contribution over (t_H^C, Ω_H^C) but the high cost industry does not. In addition, the government must also accept that combination (it prefers the combination over (t_L^C, Ω_L^C)). So, the combination of the tax rate for the low-cost industry and its contribution under separating equilibrium must be located in the area below IC_L and above IC_H and PC_L. The points in the grey colored area in figure 4.2 can constitute a separating equilibrium but no such area exists in figure 4.3. This is because $t_H^C \leq \hat{t}$ in figure 4.2 but $t_H^C > \hat{t}$ in figure 4.3.

As we can see from figure 4.2, the low cost industry must accept a higher tax rate than t_H^C (the tax rate of the high cost industry under a complete information case) to reveal its true type. Thus, the lower limit for the separating equilibrium tax rate of the low cost industry is determined by t_H^C. In addition, we can consider t_H^C as the strength of the incentive to pretend because the high cost industry has this incentive as long as the tax rate of the low cost industry is lower than t_H^C(and has a lower political contribution than Ω_H^C). However, pretending to be a high cost industry is better for a low cost industry if it has to accept a high tax rate to reveal its type. This limit that a low cost industry has is an incentive to reveal its type, and the upper limit for the separating equilibrium tax rate of the low cost industry is $\hat{t}$. Because of this upper limit, the low cost industry cannot reveal its type if the high cost industry has a strong incentive to pretend to be a low cost industry or if t_H^C is high (relative to $\hat{t}$). Otherwise, separating equilibria are likely to exist if the pretending incentive of the high cost industry is weak.

We now consider the set of parameters under which a separating equilibrium exists more concretely. When $2c\bar{e}_L \geq \bar{e}_H - \bar{e}_L$ ($\bar{e}_H$ and $\bar{e}_L$ are not so different) and $t_H^C > 0$, $\hat{t} = c(\bar{e}_L + \bar{e}_H)/[2(1+c)]$and $t_H^C = [\bar{e}_H(\alpha - c)c]/[(\alpha - 1)c + \alpha]$. Therefore, if

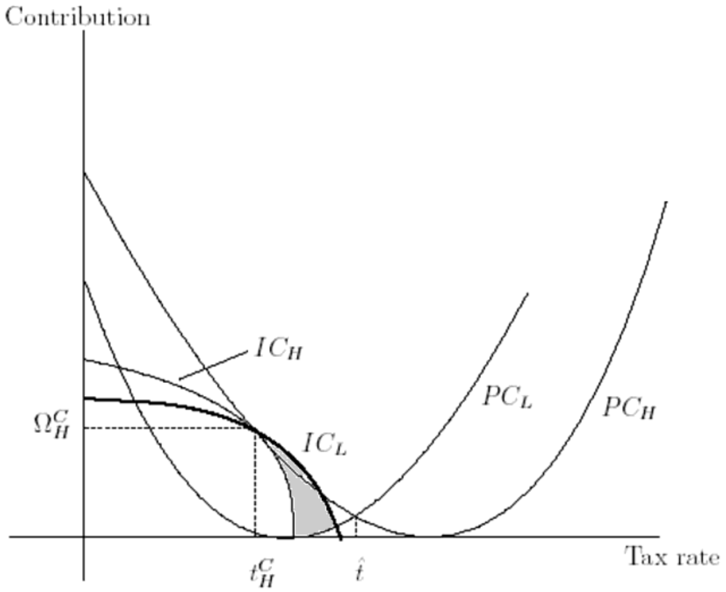
$$\frac{\bar{e}_H(\alpha - c)c}{(\alpha - 1)c + \alpha} \leq \frac{c(\bar{e}_L + \bar{e}_H)}{2(1+c)} \Leftrightarrow \frac{\alpha - c + \alpha c - 2c^2}{\alpha - c + \alpha c}\bar{e}_H \leq \bar{e}_L, \tag{4.9}$$

there exists a separating equilibrium[9]. By partially differentiating (4.9) with c and α, we get

$$\frac{\partial(LHS\ of\ (4.9))}{\partial \alpha} = \frac{2(1+c)c^2}{(\alpha - c + \alpha c)^2}\bar{e}_H > 0, \quad \frac{\partial(LHS\ of\ (4.9))}{\partial c} = \frac{-2c[(\alpha - c + \alpha c) + \alpha]}{(\alpha - c + \alpha c)^2}\bar{e}_H < 0.$$

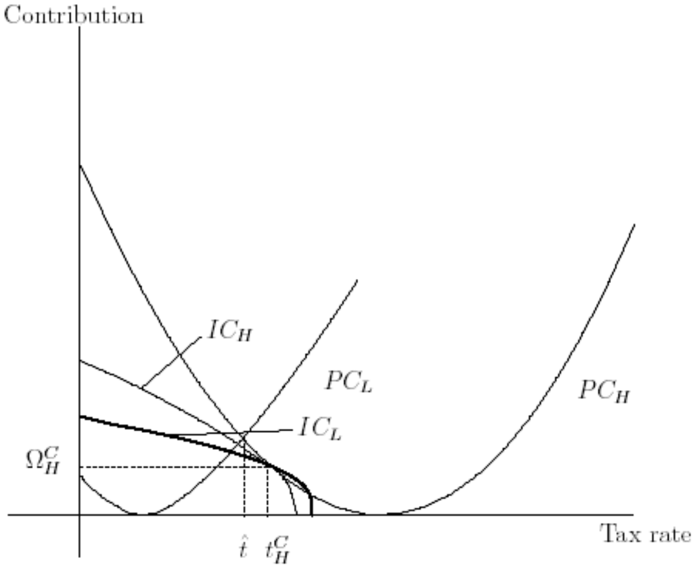
---

[9] We ignore $\alpha - c + \alpha c \leq 0$. However, in this case, $t_H^C = 0$, i.e. there always exists a separating equilibrium.



IC_i is isocost(tax+ abatement) curve of type i.
PC_i is set of (t, Ω) which satisfies PC contratint for type i with equality.

Figure 4.2: A case in which a tax separating equilibrium exists



IC_i is isocost(tax+ abatement) curve of type i.
PC_i is set of (t, Ω) which satisfies PC contratint for type i with equality.

Figure 4.3: A case in which no tax separating equilibrium exists

Hence, the smaller the difference in natural emission ($\bar{e}_H - \bar{e}_L$), the less the government cares about social welfare, and/or the steeper the slope of MAC, the more likely it is that a separating equilibrium exists[10]. This result may not be intuitive, but it is quite reasonable given that the high cost industry has less incentive to pretend to be the low cost industry, making it harder for the low cost industry to reveal its type. First, if the difference in natural emissions is smaller, then the high cost industry has more incentive to pretend to have low costs because the differences in tax rates and contributions are smaller, i.e., the benefit of pretending is smaller. Second, if the government cares less about social welfare, then the industry does not have to contribute much to set a tax rate that is different from the optimal one. Thus, the net benefit of lobbying is large and the benefit of pretending is small relative to the net benefit of lobbying. Therefore, the high cost industry has less incentive to pretend to be the low cost industry.

If the slope of the MAC is steeper, the relative cost of lobbying to the sum of the abatement and the tax is smaller, and the tax rate of the high cost industry under complete information is lower[11]. However, the socially optimal tax rates are higher, and therefore, $\hat{t}$ is high[12]. As the condition for the existence of a separating equilibrium is $t_H^C \leq \hat{t}$, the separating equilibrium is more likely to exist if the slope of the MAC is steeper.

The next proposition describes cases when private information improves social welfare.

Proposition 4.3. *Social costs under a separating equilibrium are not higher than they are under complete information when $c \leq 1$ and $\alpha \leq c$.*

Proof: The social costs of the high-cost industry under the complete information case and separating equilibrium are the same. If $\alpha \leq c$, the social cost of the low-cost industry under the complete information cases is $\bar{e}_L^2/2$ because equation (4.3) implies that the low-cost industry emits $\bar{e}_L$ in such

---

[10] One might assume that the government cannot differentiate the type of industry, whether low cost or high cost, if$\bar{e}_H = \bar{e}_L$; this is true. However, the government can differentiate the natural emission of the industry. The industry's type does not matter because what the government actually needs to set the tax rate is information on the natural emission level of the industry. Therefore, the equilibrium under the case with $\bar{e}_H = \bar{e}_L$ (equivalent to one under the complete information case) can be considered a separating equilibrium in the sense that the government has the information necessary to set the tax rate.

[11] Actually, $\frac{\partial t_H^C}{\partial c} = \bar{e}_H[(\alpha - c)c]/[(\alpha - 1)c + \alpha]^2 < 0$ if $t_H^C > 0$.

[12] Actually, $\hat{t} = (t_H^* + t_L^*)/2$, and $\frac{\partial \hat{t}}{\partial c} = \frac{(\bar{e}_H + \bar{e}_L)}{2}[(\alpha - c)c]/[1 + c]^2 < 0$ if $t_H^C > 0$.

cases. The social cost of the low-cost industry under separating equilibrium is smaller than $\bar{e}_L^2/2$ for the following reasons; (1) the low-cost industry reduces the positve amount of emissions, (2) the social cost is $c\bar{e}_L^2/2$ if the low-cost industry does not emit at all, and (3) the social cost decreases with emissions down to the socially optimal emission level, then increases until there are no emissions. *QED*

As mentioned in the proof of propostion 3, equation (4.3) states that tax rates are zero under complete information cases if $\alpha \leq c$. Therefore, the industry does not reduce its emissions at all under such cases. However, because tax rate for the low-cost industry is positive under separating equilibria, the low-cost industry reduces its emissions but it may reduce too much. The social cost in the case of excessive reduction is at most $c\bar{e}_L^2/2$ (= social cost under the no-emissions case). If $c \leq 1$, $c\bar{e}_L^2/2 < \bar{e}_L^2/2$. Therefore, private information might improve social welfare under taxes.

4.4.1.2 Pooling equilibria

Under pooling equilibria, the government cannot differentiate the types of the industries. In other words, the high cost industry has an incentive to prevent the government from differentiating it under pooling equilibira. Technically, any t, $0 \leq t \leq c\bar{e}_H$ can be a pooling equilibrium rate such that

$$\frac{1}{2}F(\bar{e}_H) \geq C_H(e_H(t)) + te_H(t) + \Omega^P(t). \tag{4.10}$$

where $\Omega^P(t) \geq \alpha E_{\bar{e}}[SC(t) - SC(t^{P*})]$ and $t^{P*} = \arg\min E_{\bar{e}}[SC(t)]$. t must satisfy $\Omega^P(t') \leq \Omega^P(t) + \alpha E_{\bar{e}}[SC(t') - SC_i(t)]$ for any $t' \neq t$. Let t_1^P and t_2^P be a smaller t and bigger t, respectively, that satisfy (4.10) with equality [13].

Although the total cost of the high-cost industry under pooling equilibria is not higher than it is under the complete information case or separating equilibria, it is unclear whether the total cost of the low-cost industry is higher than it is under separating equilibrium. Whether the total cost of the low-cost industry is higher than that under separating equilibrium depends on the pooling

---

[13] If $t_1^P < 0$, $t_1^P = 0$. If $t_2^P > c\bar{e}_H$, $t_2^P = c\bar{e}_H$.

equilibrium tax rate, i.e., the government's belief about the industry's type and parameters. In order to compare the results under separating equilibrium to those under pooling equilibrium, in the next section we focus on the equilibria where the total cost of the low-cost industry is the lowest. We can eliminate pooling equilibria by domination-based refinement of beliefs if the total cost of the low-cost industry under the best separating equilibrium is lower than the total cost under pooling equilibria. Because the total cost of the industry at the given tax rate is minimized when $\Omega^P(t) = \alpha E_{\bar{e}}[SC(t) - SC(t^{P*})]$, the tax rate at which the cost of the low cost industry is minimized[14] is

$$t_P = \arg\min_{t \geq 0}\; te_L(t) + C_L(e_L(t)) + \Omega^P(t) = \frac{1}{2}\left[\frac{(\alpha-1)c+\alpha}{c^2}t^2 + 2\left(\bar{e}_L - \frac{\alpha}{c}E[\bar{e}]\right)t + \frac{\alpha}{1+c}E[\bar{e}_L]^2\right]. \tag{4.11}$$

If $(\alpha-1)c+\alpha < 0$ or $E[\bar{e}]\alpha < c\bar{e}_L{}^{15}$, $t_P = 0$. Otherwise,

$$t_P = \begin{cases} t_2^P & if\ t_2^p < \frac{(E[\bar{e}]\alpha - c\bar{e}_L)c}{(\alpha-1)c+\alpha} \\ \frac{(E[\bar{e}]\alpha - c\bar{e}_L)c}{(\alpha-1)c+\alpha} & otherwise. \end{cases} \tag{4.12}$$

Again, this t_P will be used in the numerical example.

4.4.2 Quotas

The low-cost industry might have an incentive to pretend that it has high costs because it would benefit from a large quota. Besides, equation (4.5) implies that the high-cost industry is given a larger quota than the low-cost industry under the complete information case. On the other hand, the high-cost industry does not have an incentive to pretend that it has low costs. Therefore, in the next sub-subsection, we will describe the case in which the low-cost industry has no incentive to pretend to have high costs (separating equilibrium). Then, we will discuss the case in which the

---

[14] We assume that $(1+c)\bar{e}_L \geq E[\bar{e}]$. This assumption ensures that $t^{P*} = cE[\bar{e}]/(1+c)$ and $\frac{(E[\bar{e}]\alpha - c\bar{e}_L)c}{(\alpha-1)c+\alpha} < c\bar{e}_L$. See appendix C.1 when $(1+c)\bar{e}_L < E[\bar{e}]$.

[15] In this case, $t_P = 0$ satisfies Eq(4.10)

government cannot tell the type of the industry (pooling equilibrium).

4.4.2.1 Separating equilibria

The conditions for a separating equilibrium are the IC conditions (Eq (4.13) for the low-cost industry and (4.14)[16] for the high-cost industry),different types have to accept to reveal their type

$$C_L(q_L^C) + \Omega_L^C(q_L^C) \leq C_L(e_L(q_H)) + \Omega_H^C(q_H) \tag{4.13}$$

$$C_H(q_L^C) + \Omega_L^C(q_L^C) \geq C_H(q_H) + \Omega_H^C(q_H) \tag{4.14}$$

and PC conditions, $\Omega_i(q_i) \geq \alpha[SC_i(q_i) - SC_i(q_i^*)]$ for $i = L, H$ and $\Omega_i(q) \leq \Omega_i(q_i) + \alpha[SC_i(q) - SC_i(q_i)]$ for $q \in Q$ and $i = L, H$. q_H must be larger than q_H^*. Unlike taxes, a separating equilibrium always exists under quotas.

Proposition 4.4. *There exists a separating equilibrium under quotas.*

Proof: See appendix C.6.

The reason why a separating equilibrium always exists under quotas is as follows. The low-cost industry wants a large quota less than the high-cost industry does. Therefore, when the high-cost industry puts pressure on the government to set a larger quota, the low-cost industry has less incentive to pretend. Finally, the low-cost industry has no incentive to enlarge the quota. When a tax is implemented, the low-cost industry reveals its type by making the government set the tax rate for the low-cost industry higher than that for the high-cost industry. However, the tax rate for the high-cost industry under the complete information case is greater than that for the low-cost industry, so it is sometimes hard for the low-cost industry to make the government set the tax rate higher for the low-cost industry. This is why a tax separating equilibrium might not exist. On the other hand, the quota for the high-cost industry under complete information is larger than that for the low-cost industry, and the high-cost industry more strongly advocates for a large quota than the low-cost industry. Therefore, it is not difficult for the high-cost industry to make the government set the quota high enough that the low-cost industry will have no incentive to pretend.

---

[16] If $\bar{e}_L \leq q_H$, then this condition is the same as $C_L(q_L^C) + \Omega_L^C(q_L^C) \leq \Omega_H^C(q_H)$.

As for social welfare, the social cost under separating equilibria is not lower than that under complete information. The high-cost industry might make the government set the quota higher than it would under the complete information case in order to reveal its type[17]. The low-cost industry receives the same quota as it does in the complete information case. Thus, private information does not improve social welfare under quotas.

4.4.2.2 Pooling equilibria

Whether a separating equilibrium is achieved depends on the government's belief about the industry's type, even if it always exists. Therefore, it is possible that pooling equilibria also exist. If a pooling equilibrium exists, the low-cost industry has an incentive not to make the government differentiate its type. In this sub-subsection, we investigate the property of pooling equilibria. Technically, any q can be a pooling equilibrium such that

$$C_L(q_L^C) + \Omega_L^C(q_L^C) \geq C_L(e_L(q)) + \Omega^P(q) \qquad (4.15)$$

where $\Omega^P(q) \geq \alpha[E_{\bar{e}}[SC(q) - SC(q^*)]]$ and $q^* = \arg\min\ E_{\bar{e}}[SC(q)]$. q must satisfy $\Omega^P(q') \leq \Omega^P(q) + \alpha E_{\bar{e}}[SC(q') - SC_i(q)]$ for any $q' \neq q$. Let q_1^P and q_2^P be a smaller q and bigger q, respectively, that satisfy (4.15) with equality[18].

While the total cost of the low-cost industry under pooling equilibria is not higher than that under the complete information case or separating equilibria, it is ambiguous whether the total cost of the high-cost industry is higher than it is under separating equilibrium. Whether the total cost of the high-cost industry is higher than that under separating equilibrium depends on the pooling equilibrium quota, i.e., the government's belief about each industry's type and parameters. In order to compare the results obtained under separating equilibrium to those obtained under pooling equilibrium, in the next section we focus on the equilibrium where the total cost of the high-cost industry is the lowest. Because the total cost of the industry at the given quota is minimized when $\Omega^P(q) = \alpha E_{\bar{e}}[SC(q) - SC(q^{P*})]$, the quota at which the cost of the high-cost industry is

---

[17] Depending on parameters, the low-cost industry does not have an incentive to pretend to be a high-cost industry.

[18] If $q_1^P < 0$, $q_1^P = 0$. If $q_2^P > \bar{e}_H$, $q_2^P = \bar{e}_H$.

minimized[19] is

$$q_H^P = \arg \min_{q_1^P \le q \le q_2^P} C_H(q) + \Omega^P(q) = \frac{1}{2}\left[\{c+(1+c)\alpha\}q^2 - 2(c\bar{e}_H + \alpha c E[\bar{e}])q + c\bar{e}_H^2 + \frac{\alpha c^2}{1+c}E[\bar{e}]^2\right]. \tag{4.16}$$

From first order conditions (F.O.Cs),

$$q_H^P = \begin{cases} \frac{\alpha c E[\bar{e}]+c\bar{e}_H}{c+(1+c)\alpha} & if\ q_2^P < \frac{\alpha c E[\bar{e}]+c\bar{e}_H}{c+(1+c)\alpha} \\ q_2^P & otherwise. \end{cases} \tag{4.17}$$

4.4.3 Difference in results and incentives between tax and quota

In this subsection, we briefly and intuitively review what happens under tax and quota. In the last subsection, we showed that private information may worsen social welfare under quota but may improve it under tax. The type that has the incentive to pretend to be the other is different between tax and quota. In addition, the direction of change in the regulation level depends on which type has to reveal its type (or which type has an incentive to pretend to be the other type). These factors result in the differences under tax and quota.

The high cost industry more strongly desires lax regulation or has a higher willingness to pay than the low cost industry. Roughly speaking, the high cost industry prefers the combination of the high political contribution and lax regulation to that of the low political contribution and stringent regulation, while the low cost industry prefers the combination of low political contribution and stringent regulation. Therefore, the high cost industry can reveal its type by accepting laxer regulations and higher political contributions relative to the complete information case because benefit. If the high cost industry does so, social welfare decreases relative to the complete information case. Thus, private information worsens social welfare if the high cost industry has to reveal its type. However, the low cost industry cannot reveal its type by accepting laxer regulations and higher political contributions relative to the complete information case. If the low cost industry has to reveal its type, it must also accept stringent regulations by paying low political contributions relative to the complete

---

[19] We assume that $(1+c)\bar{e}_L \ge cE[\bar{e}]$ and $\bar{e}_L \ge q$. See appendix C.2 about other cases.

information case, which means that private information may improve social welfare when the low cost industry has to reveal its type because regulations under the complete information case are too lax due to the lobbying of the industry.

The industry prefers a high quota and low tax rate regardless of type. Under the complete information case, the quota of the high cost industry is higher than that of the low cost industry, and the tax rate of the low cost industry is lower. Under quota, the low cost industry might have an incentive to pretend to be high cost, or the high cost industry might have to reveal its type. Therefore, private information may worsen social welfare under quota. Under tax, however, private information may improve social welfare because the low cost industry may be forced to reveal its type.

Because of the two differences above, we obtain different results of the separating equilibrium under quota and tax. Private information may improve social welfare under tax but may worsen social welfare under quota. As we show in the last section, quota is generally better than tax under the complete information case due to the negative income distribution effect on the industry. In the next section, we examine how private information influences the dominance of quota over tax using numerical examples.

4.4.4 Refinements of beliefs

Which set of tax rates or quotas and political contribution schedules constitutes an equilibrium depends on the government's belief about the industry's type. The simplest refinement, the domination-based refinement of beliefs described in the appendix of Chapter 13 of Mas-colell et al. (1995), is based on the idea that reasonable beliefs should not assign a positive probability to an industry's taking an action that is strictly dominated for that industry. This refinement eliminates all separating equilibria except for the best one. However, it cannot rule out pooling equilibria, where the payoff of the low type is higher than it is in the best separating equilibrium under taxes and where the high type's payoff is higher than it is in the best separating equilibrium under quotas. Because the total costs of the industry under separating and pooling equilibria are complicated, we will examine in the next section whether pooling equilibria are eliminated by the refinement of beliefs by using numerical examples.

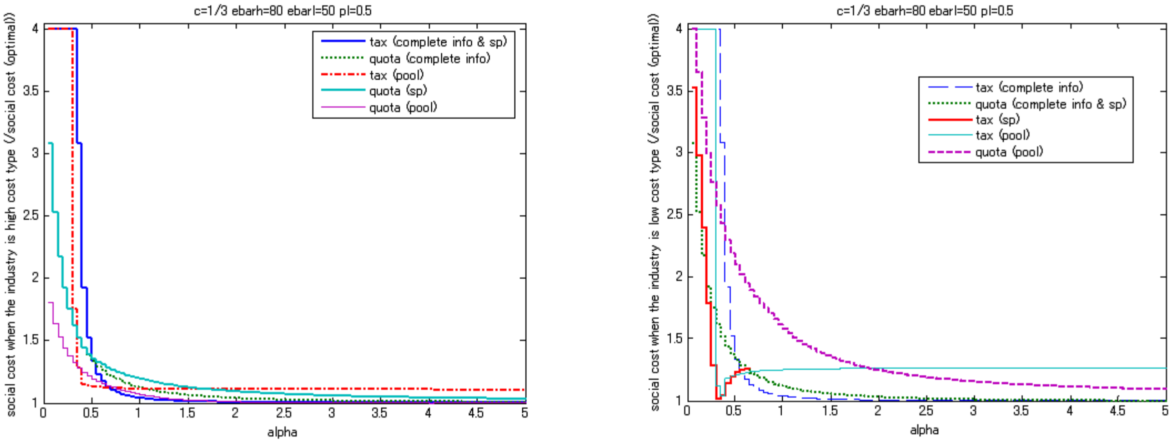
4.5 Numerical examples

Figures 4.4, 4.5, and 4.6 show social costs of the high-cost industry (figures 4.4(a), 4.5(a), and 4.6(a)) and the low-cost industry (4.4(b), 4.5(b), and 4.6(b)), the industries' costs for the high-cost industry (figures 4.4(c), 4.5(c), and 4.6(c)) and the low-cost industry (4.4(d), 4.5(d), and 4.6(d)), and political contributions (figures 4.4(e), 4.5(e), and 4.6(e)) under different equilibria with changing α when c is $1/3$, 1 or 3. The natural emission levels of the high- and low-cost industry are 80 and 50, respectively.

The political contribution of the industry under taxes and complete information is increasing with α until $\alpha = c$, the maximum value of α where the equilibrium tax rates are zero and decreasing after that. The contribution of the low-cost industry under the tax separating equilibrium is increasing with α when α is very small relative to c, decreasing until the value of α is close to c, and then increasing. In addition, no separating equilibrium exists when α is greater than the value at which the contribution of the low-cost industry under tax separating equilibrium is close to that of the high-cost industry under taxes. The political contribution under quotas with complete information is increasing with α when it is very small. Otherwise, the contribution is decreasing. The contribution of the high abatement cost industry under the separating equilibrium (the same as the complete information case) is increasing in a much wider range of α.

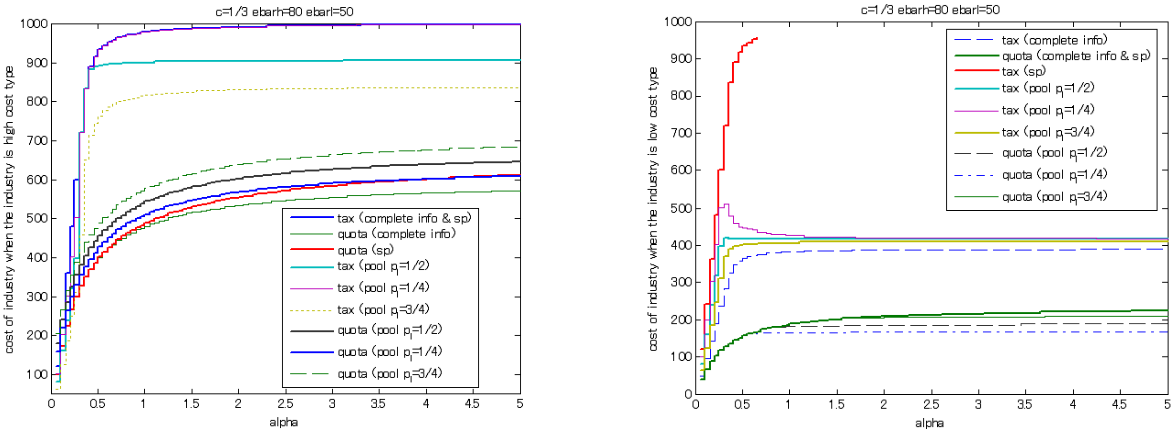
When taxes are implemented, social welfare under separating equilibrium is higher than it is under the complete information case if the industry is of low abatement cost and α is less than two times greater than c (figures 4.4(b), 4.5(b) and 4.6(b)). Hence, private information improves social welfare when the government has little concern for social welfare. The industry's cost under pooling equilibria (best for the low-cost industry) is smaller than it is under the separating equilibrium in all cases ($c = 1/3$, 1, and 3). Therefore, we cannot eliminate pooling equilibria. The industry's cost is decreasing in p_L, the prior belief that the industry is of low cost.

Under quotas, the low-cost industry has more incentive to pretend to be a high-cost industry when the slope of marginal abatement cost (MAC) is higher (c is higher) and the government cares social welfare more (i.e., α is greater). When the government cares more about social welfare, the industry has to pay a larger political contribution to get a larger quota. Thus, pretending to be a high-cost



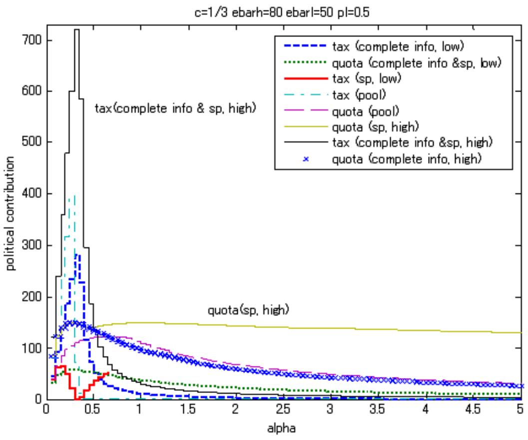
(a) Social costs of the high-cost industy

(b) Social costs of the low-cost industy



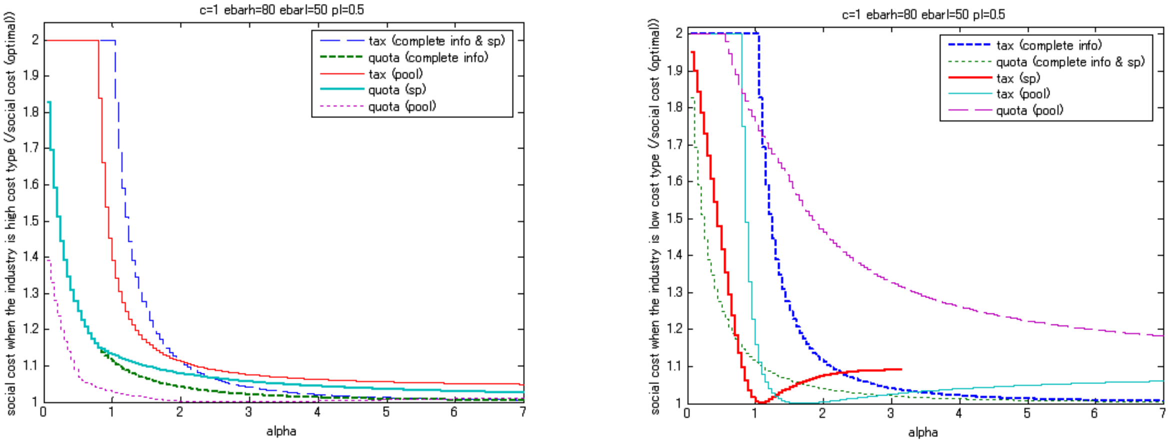
(c) High-cost industy' costs

(d) Low-cost industy' costs



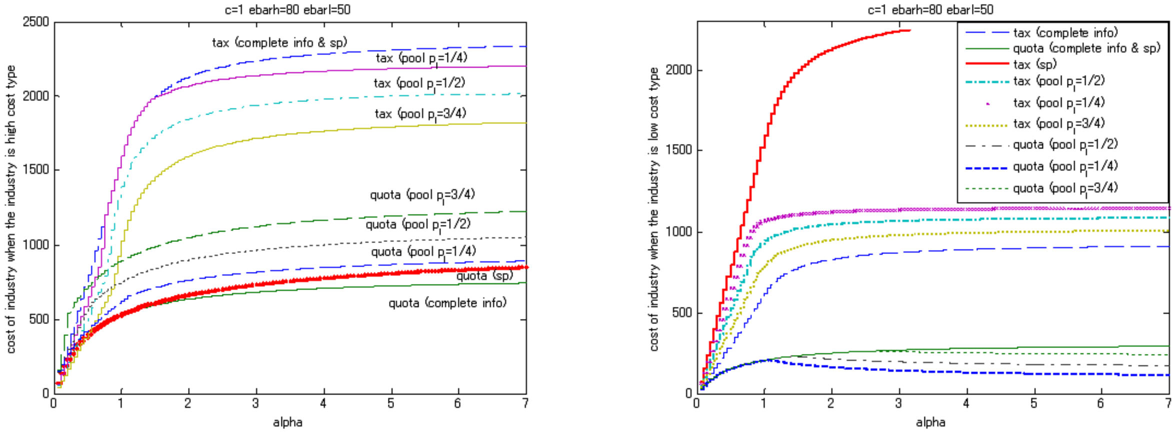
(e) Political contributions

Figure 4.4: Social costs, industries' costs, and political contributions when $c = 1/3$



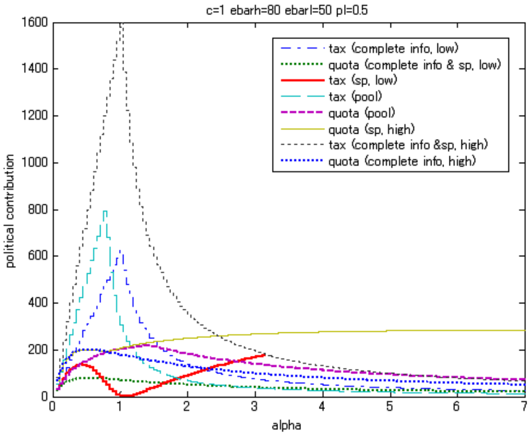
(a) Social costs of the high-cost industy

(b) Social costs of the low-cost industy



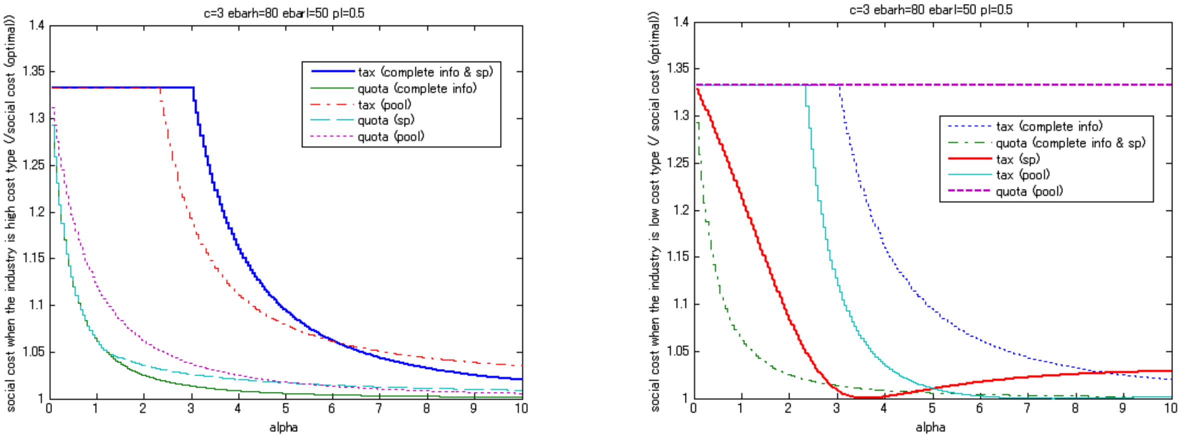
(c) High-cost industy's costs

(d) Low-cost industy's costs



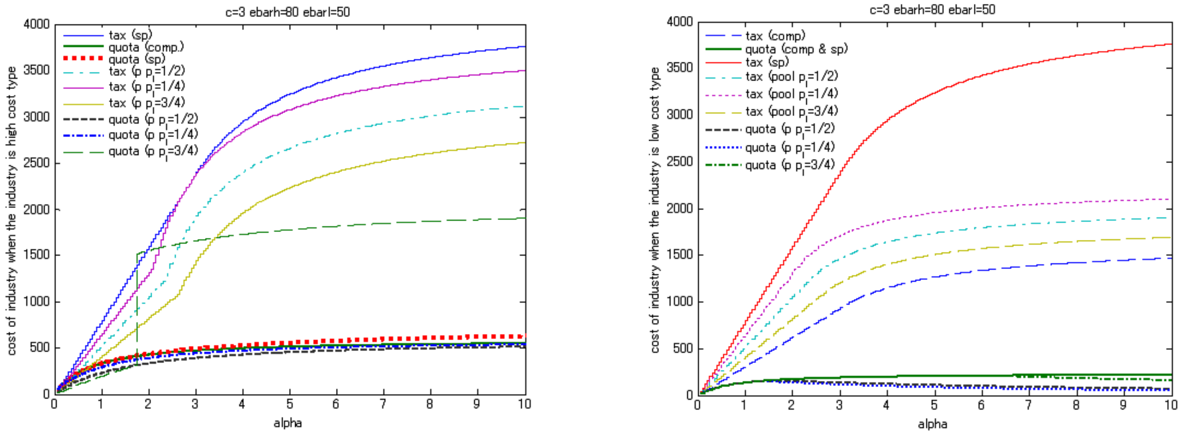
(e) Political contributions

Figure 4.5: Social costs, industries' costs, and political contributions when $c = 1$



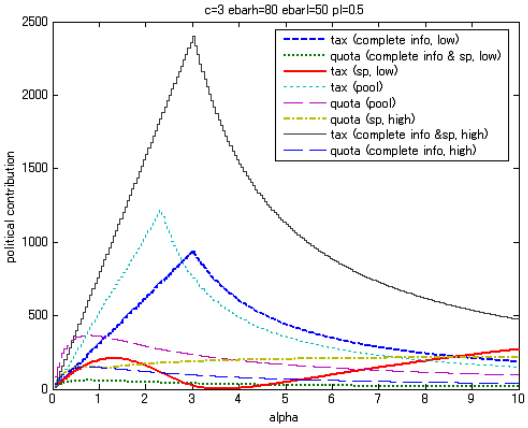
(a) Social costs of the high-cost industy

(b) Social costs of the low-cost industy



(c) High cost industy's cost

(d) Low cost industy's cost



(e) Political contributions

Figure 4.6: Social costs, industries' costs, and political contributions when $c = 3$

industry is a more attractive way for the low-cost industry to get a larger quota. When the slope of MAC is flat ($c = 1/3$ and 1), the high-cost industry's cost under the best separating equilibrium is smaller than it is under the pooling equilibrium. However, the opposite happens when the slope of MAC is steep ($c = 3$). The socially optimal quota under the pooling equilibrium is the same as that for the high-cost industry when the slope of MAC is steep and there is only a slight prior belief that the industry is of low cost. Under such a case, the industry can obtain the same quota with a smaller political contribution under a pooling equilibrium than it can under the separating equilibrium. This is why the high-cost industry's cost under the pooling equilibrium is smaller than it is under the separating equilibrium with the steep MAC curve ($c = 3$).

For comparison of the results under taxes with those under quotas (figures 4.4(c), 4.4(d), 4.5(c), 4.5(d), 4.6(c), and 4.6(d)), the industry seems to almost always prefer quotas over taxes under any type of equilibrium. This is quite intuitive because the industry must pay the tax as well as the abatement expense under taxes, whereas it pays only the abatement expense under quotas. Under the complete information case, the quota is always better than the tax if $c \geq 1$[20]. under the incomplete information case, sometimes (i.e., when the value of α is similar to that of c) the tax is better than the quota when the industry is of low cost.

Figure 4.7 shows the lowest separating equilibrium quota for the high abatement cost industry under the cases where the low abatement cost industry has different natural emission levels ($\bar{e}_L = 50$, 60, and 70) with changing α, whereas figure 4.8 shows the lowest separating equilibrium tax rate for the low abatement cost industry under the cases where the high abatement cost industry has different natural emission levels ($\bar{e}_H = 60$, 70, and 80) with changing α. The high abatement cost industry with higher natural emission has more incentive to pretend to be the low abatement cost industry (i.e., the separating equilibrium tax rate is high) because the difference in the political contributions necessary to set the same (low) tax rate between the low- and high-cost industries is larger when a high-cost industry has more natural emission. In other words, the high-cost industry with higher natural emission will get more benefits from pretending to be a low-cost industry if taxes are implemented.

If quotas are implemented, the lowest separating equilibrium quota for the low-cost industry with

---

[20] Eq (4.6) always holds if $c \geq 1$.

a lower natural emission level is smaller when α is small. This is because the quota for the high-cost industry under the complete information case is larger than the natural emission level of the low-cost industry, so the low-cost industry with a lower natural emission level has less incentive to pretend to be a high-cost industry. On the other hand, the lowest separating equilibrium quota for the low-cost industry with a lower natural emission level is smaller when α is large. In this case, the quota for the high-cost industry under the complete information case is smaller than the natural emission level of the low-cost industry. In addition, the smaller the natural emission level of the low-cost industry is, the larger the difference in the quotas under complete information between the high- and low-cost industries becomes. Therefore, the low-cost industry with a lower natural emission level has more incentive to pretend to be a high-cost industry.

4.6 Conclusion

We compare taxes with quotas and complete information cases with incomplete information cases when levels of regulation are affected by an industry lobby. We show that quotas are socially preferred over taxes when the government cares little about social benefit very much or when the slope of MAC is steeper than that of MD. However, private information might reduce the comparative disadvantage of taxes compared to quotas. This result is totally different from the result in the literature on tax versus quota. If the slope of the MAC curve is steep relative to that of the MD curve, the industry's benefits from lobbying are larger under a tax than they are under a quota, so the industry lobbies more strongly against taxes than against quotas. The impact on social welfare of this scenario is greater than that of asymmetric information. Therefore, we might be able to say that political economy issues induce greater distortion than asymmetric information does. We also show that the tax rate for the low abatement cost industry is higher than that for the high abatement cost industry if a separating equilibrium exists. This result is consistent with the demonstration byEkins and Speck (1999) that the effective tax rates for emission-intensive or manufacturing industries are much lower than the nominal tax rates for other sectors in Sweden, Denmark, and Norway.

Both pro-environmental policy groups and polluting industries lobby on environmental policies,

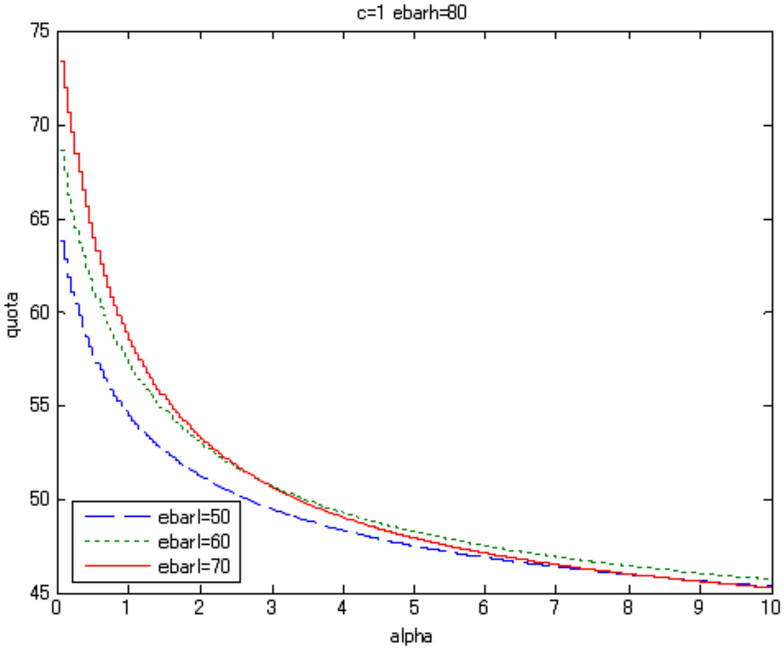


Figure 4.7: Separating equilibrium quotas for the high-cost industry with the low-cost indsutry with different levels of natural emissions($\bar{e}_H = 60$,70,and 80)

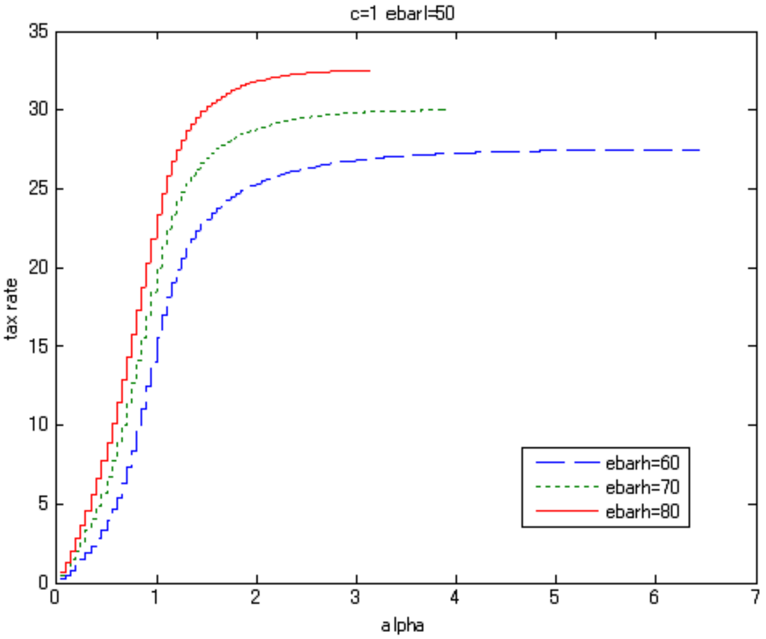


Figure 4.8: Separating equilibrium tax rate for the low-cost industry with high-cost indsutry with different levels of natural emissions($\bar{e}_H = 60$,70,and 80)

although the former's contributions are very small compared to those of the latter. Hence, extending the model of this chapter to lobbying by multiple principals with private information is one direction of future research.

Chapter 5

Conclusion

5.1 Summary of contributions

This thesis theoretically and empirically examined environmental policies, including voluntary approaches, quotas and market-based instruments. All of these policies have been effective as alternatives to traditional environmental regulation approaches such as the command and control approach.

In our theoretical analysis of voluntary emissions reduction programs, we built a model that explains why governments implement voluntary programs even though the participation rates are sometimes low. This reason has not been explained by the literature on voluntary approaches to environmental protection. Only Dawson and Segerson (2008)analyzed voluntary programs while incorporating individual firms' participation decision into their model. In their model, the program generates lower social welfare than the mandatory policy. On the other hand, our model implies that the voluntary program can generate greater social welfare and aggregate abatement than a mandatory standard if it is politically difficult to set a satisfactory standard. Therefore, regulators sometime implement a voluntary program in such a case. In this situation, the regulator sets the abatement rate of the program as high as possible, subject to the constraint that the abatement cost of participating firms is smaller than their cost under the mandatory standard. In addition to constraints on the cost of individual participating firms, no more firms will participate in the program if the aggregate abatement by the firms already participating is greater than under the mandatory standard. Our model implies that the participation rate of the program will be low if the mandatory standard is set very lax.

We empirically investigated the determinants of ISO 14001 adoption and its impacts on environmental performance in chapter 3. There are numerous studies on them, but none has addressed intra-industry spillovers of ISO 14001 adoption and environmental performance. We examined such spillovers by applying spatial econometric estimation methods to a Japanese facility dataset. We found that facilities emitting into water are more likely to adopt ISO 14001 if facilities that belong to firms with similar revenue in the same industry adopt this standard. We also found that the percentage change in emissions into air for similarly sized facilities in the same industry are correlated. The estimated spillover effects of ISO 14001 adoption between similarly sized firms emitting into water are not small. Thus, our result implies that the fact that rival firms adopt ISO 14001 is an important motive for its adoption.

In chapter 4, we examined the effects of private information about emission abatement cost on lobbying activity of a polluting industry and on social welfare under taxes and quotas. We showed that to reveal its abatement cost information, an industry with a low abatement cost makes to the government set a higher tax rate than a tax rate set under complete information that is too low in terms of the social welfare. Thus, private information might improve the social welfare under taxes, in contrast to models with a benevolent government. On the other hand, under quotas, an industry with a high abatement cost might have to make the government set a larger quota than a quota set under complete information case so that it reveals cost information. Thus, private information does not improve social welfare. In this chapter, we also compared taxes and quotas and showed that quotas are socially preferred when the slope of marginal abatement cost is steeper than that of the marginal damage or when the government weighs political contribution much more than social welfare.

5.2 Directions for further research

We conclude this thesis by noting unaddressed issues to be studied in future research. The major unaddressed issue on voluntary approaches related to environmental issues is information disclosure. Both theoretical and empirical analyses of voluntary approaches deal with voluntary actions such as improvement of environmental performance and acquirement of environmental management system certificates. However, many firms voluntarily provide self-reported information on pollution levels in their Corporate Social Responsible report, environmental reports, and/or website. Similar to the Carbon Disclosure Project (CDP), Principles for Responsible Investment (PRI), and the Investor Network on Climate Risk, we have worldwide shareholder's initiatives that require firms to disclose information on their greenhouse gas (GHG) emissions. Thus, voluntary information disclosure has been pervasive, and I believe that it is meaningful to examine determinants and impacts of voluntary information disclosure.

In particular, it is interesting to examine the relationship between firms' voluntary information disclosure and their economic and environmental performance, but we have to examine this relationship with great care. For example, firms are more likely to voluntarily disclose information on their environmental performance if their environmental performance is good. When information is available to the public, firms are subject to higher pressure to reduce emissions and actually reduce emissions. Thus, we have to control for dual causality between information disclosure and environmental performance. The credibility of self-reported information is also worth researching. For example, does third party certification on self-reported information have impacts on a firm's economic performance or its stock price? How do we make firms report truthfully without much cost? Such questions remain for to be answered by future research.

Due to the limitations of the approaches we used in this thesis, there are issues that remain to be addressed. In our analysis of voluntary pollutant reduction programs, we emphasized the political difficulty of setting mandatory policies. Heterogeneity of polluting firms is also likely to play an important role in the implementation of voluntary programs, but firms were homogeneous in our model. Introduction of the heterogeneity makes a model much more complicated, and therefore, we would have had to calculate and check many potential equilibria for the implementation of a voluntary program. In our empirical analysis, we did not determine the factors that induced the intra-industry spillovers of ISO 14001 adoption and environmental performance that we found. Therefore, identifying the causes of the intra-industry spillovers can be one direction of future research. Our model of tax versus quota analysis involved only one industry lobby group, but a model with competing lobby groups would be more realistic and might have much more relevant implications.

Bibliography

Aidt, Toke S. (1998) "Political Internalization of Economic Externalities and Environmental Policy," *Journal of Public Economics*, Vol. 69, No. 1, pp. 1-16.

Antweiler, Werner and Kathryn Harrison (2007) "Canada's Voluntary ARET Program: Limited Success Despite Industry Cosponsorship," *Journal of Policy Analysis and Management*, Vol. 26, No. 4, pp. 755-774.

Arimura, Toshi Hide, Akira Hibiki, and Hajime Katayama (2008) "Is a voluntary approach an effective environmental policy instrument?: A case for environmental management systems," *Journal of Environmental Economics and Management*, Vol. 55, No. 3, pp. 281-295.

Arimura, Toshi Hide, Nicole Darnall, and Hajime Katayama (2009) "Is ISO 14001 a Gateway to More Advanced Voluntary Action? A Case for Green Supply Chain Management," Discussion Papers dp-09-05, Resources For the Future.

Austen-Smith, David (1995) "Campaign contributions and access," *American Political Science Review*, Vol. 89, No. 3, pp. 566-581.

Austen-Smith, David and Jeffrey S. Banks (2002) "Costly signaling and cheap talk in models of political influence," *European Journal of Political Economy*, Vol. 18, No. 2, pp. 263-280.

Austen-Smith, David and John Wright (1994) "Counteractive Lobbying," *American Journal of Political Science*, Vol. 38, No. 1, pp. 25-44.

Barbier, Edward B., Richard Damania, and Daniel Leonard (2005) "Corruption, trade and resource conversion," *Journal of Environmental Economics and Management*, Vol. 50, No. 2, pp. 276-299.

Barla, Philippe (2007) "ISO 14001 certification and environmental performance in Quebec's pulp and paper industry," *Journal of Environmental Economics and Management*, Vol. 53, No. 3, pp. 291-306.

Barrett, Scott (1994) “Self-enforcing International Environmental Agreements,” *Oxford Economic Papers*, Vol. 46, pp. 878-894.

Bennedsen, Morten and Sven E. Feldmann (2006) “Informational lobbying and political contributions,” *Journal of Public Economics*, Vol. 90, No. 4-5, pp. 631-656.

Berenbeim, Ronald E. (1992) *The Corporate Ethics Practices*, New York: Conference Board.

Boyer, Marcel and Jean-Jacques Laffont (1999) “Toward a political theory of the emergence of environmental incentive regulation,” *RAND Journal of Economics*, Vol. 30, No. 1, pp. 276-299.

Brouhle, Keith, Charles Griffiths, and Ann Wolverton (2005) “The Use of Voluntary Approaches for Environmental Policymaking in the U.S,” in Croci, Edoardo ed. *The Handbook of Environmental Voluntary Agreements*, Netherlands: Springer.

Conconi, Paola (2003) “Green lobbies and transboundary pollution in large open economies,” *Journal of International Economics*, Vol. 102, No. 59, pp. 399-422.

Costello, Christopher and Larry Karp (2005) “Dynamic taxes and quotas with learning,” *Journal of Economic Dynamics and Control*, Vol. 28, No. 8, pp. 1661-1680.

Dahm, Matthias and Nicolás Porteiro (2008) “Informational lobbying under the shadow of political pressure,” *Social Choice Welfare*, Vol. 30, No. 4, pp. 531-559.

Damania, Richard (2001) “When the Weak Win: The Role of Investment in Environmental Lobbying,” *Journal of Environmental Economics and Management*, Vol. 42, No. 1, pp. 1-22.

Darnall, Nicole and Daniel Edwards (2006) “Predicting the cost of environmental management system adoption: The role of capabilities, resources and ownership structure,” *Strategic Management Journal*, Vol. 27, No. 2, pp. 301-320.

Darnall, Nicole and Alexei Pavlichev (2004) “Environmental policy tools and firm-level management in the United States.” available at: http://www.oecd.org/dataoecd/1/59/35590060.pdf.

Darnall, Nicole and Stephen Side (2008) “Assessing the performance of voluntary environmental programs: Does certification matter?” *Policy Studies Journal*, Vol. 36, No. 1, pp. 95-117.

Dasgupta, Partha, Peter Hammond, and Eric Maskin (1980) "On Imperfect Information and Optimal Pollution Control," *Review of Economic Studies*, Vol. 47, No. 5, pp. 857-860.

Dawson, Na Li and Kathleen Segerson (2008) "Voluntary Agreements with Industries: Participation Incentives with Industry-Wide Targets," *Land Economics*, Vol. 84, No. 1, pp. 97-114.

Ekins, Paul and Stefan Speck (1999) "Competitiveness and Exemptions From Environmental Taxes in Europe," *Environmental and Resource Economics*.

European Environment Agency (1997) *Environmental Agreements: Environmental Effectiveness*: Environmental Issue Series No.3 1-2.

Fredriksson, Per G. (1997) "The political economy of pollution taxes in a small open economy," *Journal of Environmental Economics and Management*, Vol. 33, No. 1, pp. 44-58.

Gamper-Rabindran, Shanti (2006) "Did the EPA's Voluntary Industrial Toxics Program Reduce Emission? A GIS Analysis of Distributional Impacts and By-media Analysis of Substitution," *Journal of Environmental Economics and Management*, Vol. 52, No. 1, pp. 391-410.

Glachant, Matthieu (2007) "Non-binding Voluntary Agreements," *Journal of Environmental Economics and Management*, Vol. 32-48, No. 1, p. 54.

Glachant, Matthieu, Marta Vicarelli, and Fr'ed'erique Vincent (2004) "Management Environnemental : Une Enquete aupres de 270 Etablissements Industriels Francais." available at: http://www.oecd.org/dataoecd/25/37/31684454.pdf.

Grossman, Gene and Elhanan Helpman (1994) "Menu auctions, resource allcation, and economic influence," *American Economic Review*, Vol. 84, No. 4, pp. 833-850.

Hansen, Lars. G. (1999) "Environmental Regulation through Voluntary Agreements," in Carraro, Carlo and François Lévêque eds. *Voluntary Approaches in Environmental Policy*, Boston: Kluwer Academic.

Hepple, Leslie W. (2004) "Bayesian Model Choice in Spatial Econometrics," in Fomby, Thomas B. and R. Carter Hill eds. *Spatial and Spatiotemporal Econometrics (Advances in Econometrics, Volume 18)*, UK: Emerald Group Publishing Limited.

Hibiki, Akira and Toshi Hide Arimura (2004) "Environmental policy tools and firm-level management in Japan." available at: http://www.oecd.org/dataoecd/26/0/31686250.pdf.

Hoel, Michael and Larry Karp (2001) "Taxes and quotas for a stock pollutant with multiplicative uncertainty," *Journal of Public Economics*, Vol. 82, No. 1, pp. 91-114.

——— (2002) "Taxes versus quotas for a stock pollutant," *Resource and Energy Economics*, Vol. 24, No. 4, pp. 367-384.

Innes, Robert and Abdoul G. Sam (2008) "Voluntary Pollution Reductions and the Enforcement of Environmental Law: An Empirical Study of the 33/50 Program," *Journal of Law and Economics*, Vol. 51, No. 2, pp. 271-296.

Kaplow, Louis and Steven Shavell (2002) "On the Superiority of Corrective Taxes to Quantity Regulation," *American Law and Economics Review*, Vol. 4, No. 1, pp. 1-17.

Karp, Larry S. and Jiangfeng Zhang (2002) "Controlling a Stock Pollutant with Endogenous Investment and Asymmetric Information," working papers, Department of Agricultural and Resource Economics, University of California, Berkeley.

Kerekes, Sándor, Gábor Harangozó, Patrícia Németh, and Á. Nemcsicsné Zsóka (2004) "Environmental policy tools and firm-level management in Hungary." available at: http://www.oecd.org/dataoecd/26/0/31686250.pdf.

Khanna, Madhu (2007) "The U.S. 33/50 Voluntary Program: Its Design and Effectiveness," in Morgenstern, Richard D. and William Aaron Pizer eds. *Reality Check the Nature and Performance of Voluntary Environmental programs in the United States, Europe, and Japan*, Washington DC: Resources for the Future.

King, Andrew A., Michael J. Lenox, and Ann Terlaak (2005) “The strategic use of decentralized institutions: Exploring certification with the ISO 14001 management standard,” *Journal of Environmental Economics and Management*, Vol. 48, No. 6, pp. 1091-1106.

LeSage, James and R. Kelly Pace (2009) *Introduction to Spatial Econometrics*, Boca Raton, FL: CRC Press.

Lewis, Tracy R. (1996) “Protecting the Environment When Costs and Benefits are Privately Known,” *RAND journal of Economics*, Vol. 27, No. 4, pp. 819-847.

Lohmann, Suzanne (1995) “Information, Access, and Contributions: A Signaling Model of Lobbying,” *Public Choice*, Vol. 85, No. 3-4, pp. 267-284.

Lutz, Stefan, Thomas P. Lyon, and John W. Maxwell (2000) “Quality leadership when regulatory standards are forthcoming,” *Journal of Industrial Economics*, Vol. 48, No. 3, pp. 331-348.

Lyon, Thomas P. and John W. Maxwell (2002) “Voluntary Approaches to Environmental Regulation: A Survey,” in Franzini, Maurizio and Antonio Nicita eds. *Economic Institutions and Environmental Policy: Post Present and Future*, Hampshire, UK: Ashgate Publishing Ltd.

——— (2003) “Self-regulation, Taxation and Public Voluntary Environmental Agreements,” *Journal of Public Economics*, Vol. 87, No. 3-4, pp. 1453-1486.

——— (2004) *Corporate Environmentalism and Public Policy*, Cambridge, UK: Cambridge University Press.

Manzini, Paola and Marco Mariotti (2003) “A Bargaining Model of voluntary Environmental Agreements,” *Journal of Public Economics*, Vol. 87, No. 12, pp. 2725-2736.

Mas-colell, Andrew, Michael D. Whinston, and Jerry R. Green (1995) *Microeconomic Theory*, New York: Oxford University Press.

Maskin, Eric and Jean Tirole (1992) “The principal-agent relationship with an informed principal, II: Common values,” *Econometria*, Vol. 60, No. 1, pp. 1-42.

Maxwell, John W., Thomas P. Lyon, and Steven C. Hackett (2000) "Self-Regulation and Social Welfare: The Political Economy of Corporate Environmentalism," *Journal of Law and Economics*, Vol. 43, No. 3, pp. 583-618.

Melnyk, Steven A., Robert P. Sroufe, and Roger Calantone (2003) "Assessing the impact of environmental management systems on corporate and environmental performance," *Journal of Operations Management*, Vol. 21, No. 3, pp. 329-351.

Moledina, Amyaz A., Jay S. Coggins, Stephen Polasky, and Christopher Costello (2003) "Dynamic Environmental Policy with Strategic Firms: Prices versus Quantities," *Journal of Environmental Economics and Management*, Vol. 45, No. 2, pp. 356-376.

Moore, Megan (2007) "Energy & Environmental Giving in the States," research report, The National Institute on Money in State Politics.

Nakamura, Masao, Takuya Takahashi, and Ilan Vertinsky (2001) "Why Japanese Firms Choose to Certify: A Study of Managerial Responses to Environmental Issues," *Journal of Environmental Economics and Management*, Vol. 42, No. 1, pp. 23-52.

Newell, Richard G. and William A. Pizer (2003) "Regulating Stock Externalities under Uncertainty," *Journal of Environmental Economics and Management*, Vol. 45, No. 2, pp. 416-432.

Nishitani, Kimitaka (2009) "An empirical study of the initial adoption of ISO 14001 in Japanese manufacturing firms," *Ecological Economics*, Vol. 60, No. 3, pp. 669-679.

OECD (1999) *Voluntary Approaches for Enivronmental Policy: An Assessment*, Paris, France.

Potoski, Matthew and Aseem Prakash (2005a) "Covenants with weak swords: ISO 14001 and facilities' environmental performance," *Journal of Policy Analysis and Management*, Vol. 24, No. 4, pp. 745-769.

——— (2005b) "Green clubs and voluntary governance: ISO 14001 firms' regulatory compliance," *American Journal of Political Science*, Vol. 49, No. 2, pp. 235-248.

Quirion, Philippe (2004) “Regulating Stock Externalities under Uncertainty,” *Environmental and Resource Economics*, Vol. 29, No. 3, pp. 337-359.

Rennings, Klaus, Manuel Frondel, Jens Horbach, and Till Requate (2004) “Environmental policy tools and firm-level management in Germany.” available at: http://www.oecd.org/dataoecd/26/17/31685533.pdf.

Segerson, Kathleen and Thomas J. Miceli (1998) “Voluntary Environmental Agreements: Good or Bad News for Enviornmental Quality,” *Journal of Environmental Economics and Management*, Vol. 36, No. 2, pp. 109-130.

——— (1999) “Voluntary Approaches to environmental protection: The Role of Legislative threats,” in Carraro, Carlo and François Lévêque eds. *Voluntary Approaches in Environmental Policy*, Boston: Kluwer Academic.

Teikoku Data Bank (2003a) *Company Yearbook East Japan*: Teikoku Data Bank.

——— (2003b) *Company Yearbook West Japan*: Teikoku Data Bank.

——— (2004a) *Company Yearbook East Japan*: Teikoku Data Bank.

——— (2004b) *Company Yearbook West Japan*: Teikoku Data Bank.

——— (2005a) *Company Yearbook East Japan*: Teikoku Data Bank.

——— (2005b) *Company Yearbook West Japan*: Teikoku Data Bank.

US Environmental Protection Agency (1999) *33/50 Program: The Final Record.*, Washington DC: Office of Pollution Prevention and Toxics (7408) U.S. EPA.

——— (2010) *Risk-screening environmental indicators user’s manual, Technical Appendix A, Listing of all toxicity weights for TRI chemicals and chemical categories*. available at: http://www.epa.gov/opptintr/rsei/pubs/technical_appendixa_toxicity.pdf.

Weitzman, Martin L. (1974) “Prices vs. Quantities,” *Review of Economic Studies*, Vol. 41, No. 4, pp. 477-491.

Welch, Eric and Akira Hibiki (2003) “An Institutional Framework for Analysis of Voluntary Policy: The Case of Voluntary Pollution Prevention Agreements in Kita Kyushu, Japan,” *Journal of Environmental Planning and Management*, Vol. 46, No. 4, pp. 523-554.

Welch, Eric W., Yasuhumi Mori, and Midori Aoyagi-Usui (2002) “Voluntary adoption of ISO 14001 in Japan: mechanisms, stages and effects,” *Business Strategy and the Environment*, Vol. 11, No. 1, pp. 43-62.

Ytterhus, Bjarne E. (2004) “Environmental policy tools and firm-level management in Norway.” available at: http://www.oecd.org/dataoecd/25/37/31684454.pdf.

Yu, Zhihao (2005) “Environmental Protection: A Theory of Direct and Indirect Competition for Political Influence,” *Review of Economic Studies*, Vol. 72, No. 1, pp. 269-286.

Appendices

Appendix A: Appendix for Chapter 2

A.1 A set of abatement rate and the number of participating firms that generates the highest aggregate abatement if N_P^V is not an integer

In this appendix, we show the set of abatement rates and the number of participating firms that generate the highest aggregate abatement when N_P^V is not an integer. Let N'_P be an integer such that $(N'_P + \varepsilon - 1)\alpha^V = N\alpha^L$ where $\alpha^L = 1 - \frac{c}{d\bar{e}_i N} - \frac{c\lambda}{d\bar{e}_i N^2(1-\lambda)}$, $\alpha^V = 1 - \frac{c}{d\bar{e}_i N} - \frac{c\lambda}{d\bar{e}_i N^2(1-\lambda)} + \frac{c\lambda}{2d\bar{e}_i N^3(1-\lambda)} = \alpha^L + \frac{c\lambda}{2d\bar{e}_i N^3(1-\lambda)}$, and $0 < \varepsilon < 1$. This N'_P is the greatest integer that satisfies (2.10) when the abatement rate satisfies (2.8) with equality. Let α'^V be the abatement rate such that $N'_P\alpha'^V = N\alpha^L$. We can then describe the equilibrium abatement rate and the number of participating firms.

Proposition A.1. *α'^V and $N'_P + 1$ are the equilibrium abatement rate and the number of participating firms that generate the highest aggregate abatement if*

$$2\{\varepsilon N - (1-\varepsilon)^2\}\{d(1-\lambda)\bar{e}_i N^3 - cN^2(1-\lambda) - Nc\lambda\} - (1-\varepsilon)^2 c\lambda > 0. \tag{1}$$

Otherwise, α^V and N'_P are the equilibrium abatement rate and the number of participating firms.

Proof: From the argument given prior to Proposition 1, it is obvious that α^V and N'_P or α'^V and $N'_P + 1$ are the equilibrium abatement rate and the number of participating firms. From simple computation, we get

$$\begin{aligned}\alpha^V N'_P &= \alpha^V(N'_P + \varepsilon - 1) + \alpha^V(1-\varepsilon)\\ &= \alpha^L N + \alpha^V(1-\varepsilon).\end{aligned} \tag{2}$$

and

$$\begin{aligned}\alpha'^V(N'_P+1) &= \alpha'^V N'_P + \alpha'^V \\ &= \alpha^L N + \alpha^L N/N'_P. \qquad (3)\end{aligned}$$

because $(N'_P+\varepsilon-1)\alpha^V = N\alpha^L$, and $N'_P\alpha'^V = N\alpha^L$. From Lemma 1, the regulator will choose α'^V if $\alpha^V N'_P < \alpha'^V(N'_P+1)$. Therefore, the regulator will choose α'^V if

$$\begin{aligned}\alpha'^V(N'_P+1) - \alpha^V N'_P &= \alpha^L N/N'_P - \alpha^V(1-\varepsilon) \\ &= \frac{1}{N'_P}[N\alpha^L - (N\frac{\alpha^L}{\alpha^V}+1-\varepsilon)\alpha^V(1-\varepsilon)] \\ &= \frac{1}{N'_P}[\varepsilon N\alpha^L - (1-\varepsilon)^2\alpha^V] \\ &= \frac{1}{N'_P}\left\{[\varepsilon N-(1-\varepsilon)^2](1-\frac{c}{d\bar{e}_iN}-\frac{c\lambda}{d(1-\lambda)\bar{e}_iN^2})-(1-\varepsilon)^2\frac{c\lambda}{2d(1-\lambda)\bar{e}_iN^3}\right\} \\ &= \frac{2\{\varepsilon N-(1-\varepsilon)^2\}\{d(1-\lambda)\bar{e}_iN^3-cN^2(1-\lambda)-Nc\lambda\}-(1-\varepsilon)^2c\lambda}{2N'_pd(1-\lambda)\bar{e}_iN^3} \\ &> 0.\end{aligned}$$

Therefore, $\alpha'^V(N'_P+1) > \alpha^V N'_P$ if (1) holds. *QED*

A.2 Complete proof of the statement that there is equilibrium only when less than ($N_P^V - 1$) firms participate in the VP and the mandatory policy is implemented (the second part of proposition 2.2)

The mandatory policy generates a (weakly) higher aggregate abatement if the abatement rate is α^V and if the number of participating firms is less than (N_P^V-1) because $(N_P^V-1)\alpha^V\bar{e}_i = N\alpha^L\bar{e}_i$. Therefore, the mandatory policy is implemented under such a case.

Suppose that N_P (less than N_P^V-1) firms participate in the VP. As long as $N_P < N_P^V$, the mandatory policy is implemented. Even if the VP is implemented, participating firms get the same payoff as they do when the mandatory policy is implemented. Therefore, if $N_P < N_P^V$, then all participating firms do not have to change their decision (from “participate” to “not participate”) In addition, all “non-participating” firms do not have to change their decision (from “not participate” to “participate”). *QED*

A.3 Proof of Proposition 2.3, 2.5 and 2.6

We give proof of Propositions 2.3, 2.5, and 2.6 in the Appendix. We show the effects of changes in each parameter on the abatement rate of the VP and the mandatory policy, the participation rate of the VP, and the aggregate abatement of the two programs. We have$\alpha^L = 1 - \frac{c}{d\bar{e}_i N} - \frac{c\lambda}{d\bar{e}_i N^2(1-\lambda)}$, $\alpha^V = 1 - \frac{c}{d\bar{e}_i N} - \frac{c\lambda}{d\bar{e}_i N^2(1-\lambda)} + \frac{c\lambda}{2d\bar{e}_i N^3(1-\lambda)}$, $N_P^V/N = \frac{\alpha^L}{\alpha^V} + 1 = 1 + 1/N - \frac{c\lambda}{2d\bar{e}_i N^3(1-\lambda) - 2N^2(1-\lambda)c - 2Nc\lambda + c\lambda}$, and $N_P^V \alpha^V \bar{e}_i = N\alpha^L \bar{e}_i + \alpha^V \bar{e}_i$. Let $A = 2d\bar{e}_i N^3(1-\lambda)$ and $B = \frac{c\lambda}{2dN^3(1-\lambda)}$. then,

($\bar{e}_i$) $0 \leq \frac{\partial \alpha^V}{\partial \bar{e}_i} = \frac{c}{dN\bar{e}_i^2} + \frac{c\lambda 2dN^3(1-\lambda)}{A^2}(2N-1) \leq \frac{c}{dN\bar{e}_i^2} + \frac{c\lambda 2dN^3(1-\lambda)}{A^2}2N = \frac{\partial \alpha^L}{\partial \bar{e}_i}$. $\frac{\partial (N_P^V/N)}{\partial \bar{e}_i} = \left(\alpha^V\right)^{-2}(\alpha^V \frac{\partial \alpha^L}{\partial \bar{e}_i} - \alpha^L \frac{\partial \alpha^V}{\partial \bar{e}_i}) \geq 0$. $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial \bar{e}_i} = \frac{\partial N\alpha^L \bar{e}_i}{\partial \bar{e}_i} + \frac{\partial \alpha^V}{\partial \bar{e}_i}\bar{e}_i \geq \frac{\partial N\alpha^L \bar{e}_i}{\partial \bar{e}_i}$.

(N) $0 \leq \frac{\partial \alpha^V}{\partial N} = \frac{c}{dN^2\bar{e}_i} + \frac{c\lambda 2dN^2\bar{e}_i(1-\lambda)}{A^2}(4N-3) = \frac{c}{dN^2\bar{e}_i} + \frac{c\lambda 6dN^2\bar{e}_i(1-\lambda)}{A^2}(2N-1) - \frac{2c\lambda}{A} \leq \frac{c}{dN^2\bar{e}_i} + \frac{c\lambda 6dN^2\bar{e}_i(1-\lambda)}{A^2}2N = \frac{\partial \alpha^L}{\partial N}$. $\frac{\partial N_P^V}{\partial N} = \frac{\alpha^L}{\alpha^V} + \left(\alpha^V\right)^{-2} N[\alpha^V \frac{\partial \alpha^L}{\partial N} - \alpha^L \frac{\partial \alpha^V}{\partial N}] = 1 - \left(\alpha^V\right)^{-2}[NB\frac{\partial \alpha^L}{\partial N} + N\alpha^L \frac{\partial B}{\partial N} - B(\alpha^L + B)] \geq 1$ because $N\frac{\partial B}{\partial N} = 3B$ and $\frac{\partial \alpha^L}{\partial N} = \frac{d}{dN^2\bar{e}_i} + 4B$. If $\frac{\partial N_P^V}{\partial N} \geq 1$, then $\frac{\partial (N_P^V/N)}{\partial N} \geq 0$. Therefore, $\frac{\partial (N_P^V/N)}{\partial N} \geq 0$. $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial N} = \frac{\partial N\alpha^L \bar{e}_i}{\partial N} + \frac{\partial \alpha^V}{\partial N}\bar{e}_i \geq \frac{\partial N\alpha^L \bar{e}_i}{\partial N}$.

(d) $0 \leq \frac{\partial \alpha^V}{\partial d} = \frac{c}{d^2N\bar{e}_i} + \frac{c\lambda 2\bar{e}_i N^3(1-\lambda)}{A^2}(2N-1) \leq \frac{c}{d^2N\bar{e}_i} + \frac{c\lambda 2\bar{e}_i N^3(1-\lambda)}{A^2}2N = \frac{\partial \alpha^L}{\partial d}$. $\frac{\partial (N_P^V/N)}{\partial d} = \left(\alpha^V\right)^{-2}(\alpha^V \frac{\partial \alpha^L}{\partial d} - \alpha^L \frac{\partial \alpha^V}{\partial d}) \geq 0$. $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial d} = \frac{\partial N\alpha^L \bar{e}_i}{\partial d} + \frac{\partial \alpha^V}{\partial d}\bar{e}_i \geq \frac{\partial N\alpha^L \bar{e}_i}{\partial d}$.

(c) $0 \geq \frac{\partial \alpha^V}{\partial c} = -\frac{1}{dN\bar{e}_i} - \frac{\lambda 2d\bar{e}_i N^3(1-\lambda)}{A}(2N-1) \geq -\frac{1}{dN\bar{e}_i^2} - \frac{\lambda 2d\bar{e}_i N^3(1-\lambda)}{A}2N = \frac{\partial \alpha^L}{\partial c}$. $\frac{\partial (N_P^V/N)}{\partial c} = \left(\alpha^V\right)^{-2}(\alpha^V \frac{\partial \alpha^L}{\partial c} - \alpha^L \frac{\partial \alpha^V}{\partial c}) \leq 0$. $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial c} = \frac{\partial N\alpha^L \bar{e}_i}{\partial c} + \frac{\partial \alpha^V}{\partial c}\bar{e}_i \leq \frac{\partial N\alpha^L \bar{e}_i}{\partial c}$.

(λ) $0 \geq \frac{\partial \alpha^V}{\partial \lambda} = \frac{-2cd\bar{e}_i N^3}{A^2}(2N-1) \geq \frac{-2cd\bar{e}_i N^3}{A^2}2N = \frac{\partial \alpha^L}{\partial \lambda}$. $\frac{\partial (N_P^V/N)}{\partial \lambda} = \left(\alpha^V\right)^{-2}(\alpha^V \frac{\partial \alpha^L}{\partial \lambda} - \alpha^L \frac{\partial \alpha^V}{\partial \lambda}) \leq 0$. $\frac{\partial N_P^V \alpha^V \bar{e}_i}{\partial \lambda} = \frac{\partial N\alpha^L \bar{e}_i}{\partial \lambda} + \frac{\partial \alpha^V}{\partial \lambda}\bar{e}_i \leq \frac{\partial N\alpha^L \bar{e}_i}{\partial \lambda}$. *QED*

A.4 Proof of Proposition 2.4

When $N\bar{e}_i = N'\bar{e}_i'$, $N > N'$, $c = c'$, $d = d'$, and $\lambda = \lambda'$, $\alpha^* = 1 - \frac{c}{dN\bar{e}_i} = 1 - \frac{c}{dN'\bar{e}_i'} = \alpha^{*\prime}$. From a simple calculation, we get $N_P^V/N = 1 - \frac{c\lambda}{2dN^3\bar{e}_i(1-\lambda)} + 1/N$ and $N_P^{V\prime}/N' = 1 - \frac{c\lambda}{2dN^{3\prime}\bar{e}_i(1-\lambda)} + 1/N'$.

Let $A = \frac{c\lambda}{2dN\bar{e}_i(1-\lambda)}(= \frac{c\lambda}{2dN'\bar{e}'_i(1-\lambda)})$. Then, $N_P^V/N - N_P^{V\prime}/N' = A(1/N^2 - 1/N'^2) + 1/N - 1/N' = \frac{1}{N^2N'^2}(A(N'^2 - N^2) - NN'(N' - N)) = \frac{N'-N}{N^2N'^2}(A(N'+N) - NN')$. As $N' < N$, $N_P^V/N - N_P^{V\prime}/N' > 0$ if $A(N'+N) < NN'$. Because we focus on a case with $\alpha^{L\prime} = 1 - \frac{c}{dN\bar{e}_i} - \frac{c\lambda}{dN^2\bar{e}_i(1-\lambda)} = 1 - \frac{c}{dN\bar{e}_i} - 2A/N' > 0$ as we mentioned in footnote 6 of Chapter 2, $2A = \frac{c\lambda}{dN'\bar{e}'_i(1-\lambda)} < N'$ must hold. Therefore, $A(N'+N) < \frac{1}{2}N'(N'+N) < NN'$ and $N_P^V/N - N_P^{V\prime}/N' > 0$ because $N > N'$.

Because $\alpha^V = 1 - \frac{c}{dN\bar{e}_i} - \frac{c\lambda}{dN^2\bar{e}_i(1-\lambda)} + \frac{c\lambda}{2dN^3\bar{e}_i(1-\lambda)} = 1 - \frac{c}{dN\bar{e}_i} - 2A/N + A/N^2$ and $\alpha^{V\prime} = 1 - \frac{c}{dN'\bar{e}'_i} - \frac{c\lambda}{dN'^2\bar{e}'_i(1-\lambda)} + \frac{c\lambda}{2dN'^3\bar{e}'_i(1-\lambda)} = 1 - \frac{c}{dN'\bar{e}_i} - 2A/N' + A/N'^2$, $\alpha^V - \alpha^{V\prime} = -2A/N + A/N^2 + 2A/N' - A/N'^2 = 2A/N^2N'^2(N'^2 - N^2 - 2N'N(N' - N)) = 2A(N' - N)/N^2N'^2(N' + N - 2N'N)$. Because $N > N' \geq 2$, $(N' - N) < 0$ and $(N' + N - 2N'N) < 0$. Thus, $\alpha^V - \alpha^{V\prime} > 0$. *QED*

A.5 Case without free ride in lobbying

If we assume that all firms cooperate in lobbying and that no firms free-ride, then the lobby group minimizes $C(\Omega) = \sum_{i=1}^{N} c\alpha(\Omega)\bar{e}_i + \Omega$ subject to $\frac{\partial L}{\partial \Omega} = 0$ and $L(\alpha, \Omega) = L(\alpha^*, 0)$. From the F.O.C and $L(\alpha, \Omega) = L(\alpha^*, 0)$, we obtain $\alpha^L = 1 - \frac{c}{dN\bar{e}_i(1-\lambda)}$ and $\Omega = \frac{c^2\lambda}{2d(1-\lambda)}$. If we assume the political contribution is paid equally by all firms, then $\alpha^V = 1 - \frac{c(2-\lambda)}{2dN\bar{e}_i(1-\lambda)}$ and $N_P^V/N = \frac{\alpha^L}{\alpha^V} + 1/N$. Under a case with free-riding in lobbying, we obtain the following result, which is completely different from that without free-riding.

Proposition A.2. *If $N\bar{e}_i = N'\bar{e}'_i$, $N > N'$, $c = c'$, $d = d'$, and $\lambda = \lambda'$, then $\alpha^V = \alpha^{V\prime}$, $N_P^V/N < N_P^{V\prime}/N'$, and $N_P^V\alpha^V\bar{e}_i < N_P^{V\prime}\alpha^{V\prime}\bar{e}'_i$.*

Proof: Because $N\bar{e}_i = N'\bar{e}'_i$, $\alpha^V = 1 - \frac{c(2-\lambda)}{2dN\bar{e}_i(1-\lambda)} = 1 - \frac{c(2-\lambda)}{2dN'\bar{e}'_i(1-\lambda)} = \alpha^{V\prime}$. $N_P^V/N = \frac{\alpha^L}{\alpha^V} + 1/N < \frac{\alpha^{L\prime}}{\alpha^{V\prime}} + 1/N' = N_P^{V\prime}/N'$ as $\alpha^L = \alpha^{L\prime}, \alpha^V = \alpha^{V\prime}$and $N < N'$. $N_P^V\alpha^V\bar{e}_i = N\alpha^L\bar{e}_i + \alpha^V\bar{e}_i < N'\alpha^{L\prime}\bar{e}'_i + \alpha^{V\prime}\bar{e}'_i = N_P^{V\prime}\alpha^{V\prime}\bar{e}'_i$ as $\bar{e}_i < \bar{e}'_i$, $\alpha^L = \alpha^{L\prime}, \alpha^V = \alpha^{V\prime}$ and $N\bar{e}_i = N'\bar{e}'_i$. *QED*

Proposition A.3. *(1)$\frac{\partial(N_P^V/N)}{\partial \bar{e}_i} \geq 0$, $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial \bar{e}_i} \geq 0$, $\frac{\partial \alpha^V}{\partial \bar{e}_i} \geq 0$, (2) $\frac{\partial(N_P^V/N)}{\partial N} \geq 0$, $\frac{\partial \alpha^V}{\partial N} \geq 0$, $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial N} \geq 0$, (3) $\frac{\partial(N_P^V/N)}{\partial d} \geq 0$, $\frac{\partial \alpha^V}{\partial d} \geq 0$, $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial d} \geq 0$, (4) $\frac{\partial(N_P^V/N)}{\partial c} \leq 0$, $\frac{\partial \alpha^V}{\partial c} \leq 0$, $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial c} \leq 0$, and (5) $\frac{\partial(N_P^V/N)}{\partial \lambda} \leq 0$, $\frac{\partial \alpha^V}{\partial \lambda} \leq 0$, $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial \lambda} \leq 0$.*

Proof: We have $\alpha^L = 1 - \frac{c}{dN\bar{e}_i(1-\lambda)}$, $\alpha^V = 1 - \frac{c(2-\lambda)}{2dN(1-\lambda)\bar{e}_i}$, $N_P^V/N = \frac{\alpha^L}{\alpha^V} + 1/N$ and $N_P^V = [1 - \frac{c\lambda}{2dN(1-\lambda)\bar{e}_i - c(2-\lambda)}]N + 1$. From a simple calculation, we find

$$\Delta SC = \alpha^V \bar{e}_i(\frac{1}{2}d\alpha^V \bar{e}_i - \frac{c\lambda}{N(1-\lambda)}) = \alpha^V \bar{e}_i[\frac{1}{2}d\bar{e}_i - \frac{c}{2N} - \frac{c\lambda}{2N^2(1-\lambda)} + \frac{c\lambda}{4N^3(1-\lambda)} - \frac{c\lambda}{N(1-\lambda)}] > 0.$$

Hence,

($\bar{e}_i$) $0 \leq \frac{\partial\alpha^V}{\partial\bar{e}_i} = \frac{c(2-\lambda)}{2dN(1-\lambda)\bar{e}_i^2} \leq \frac{2c}{2dN(1-\lambda)\bar{e}_i^2} = \frac{\partial\alpha^L}{\partial\bar{e}_i}$. $\frac{\partial(N_P^V/N)}{\partial\bar{e}_i} = \left(\alpha^V\right)^{-2}(\alpha^V\frac{\partial\alpha^L}{\partial\bar{e}_i} - \alpha^L\frac{\partial\alpha^V}{\partial\bar{e}_i}) \geq 0$. $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial\bar{e}_i} = \frac{\partial N\alpha^L\bar{e}_i}{\partial\bar{e}_i} + \frac{\partial\alpha^V}{\partial\bar{e}_i}\bar{e}_i \geq \frac{\partial N\alpha^L\bar{e}_i}{\partial\bar{e}_i}$.

(N) $0 \leq \frac{\partial\alpha^V}{\partial N} = \frac{c(2-\lambda)}{2dN^2(1-\lambda)\bar{e}_i} \leq \frac{2c}{2dN^2(1-\lambda)\bar{e}_i} = \frac{\partial\alpha^L}{\partial N}$. $\frac{\partial N_P^V}{\partial N} = 1 + \frac{c^2\lambda(2-\lambda)}{[2dN(1-\lambda)\bar{e}_i - c(2-\lambda)]^2} \geq 1$. If $\frac{\partial N_P^V}{\partial N} \geq 1$, then $\frac{\partial(N_P^V/N)}{\partial N} \geq 0$. Therefore, $\frac{\partial(N_P^V/N)}{\partial N} \geq 0$. $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial N} = \frac{\partial N\alpha^L\bar{e}_i}{\partial N} + \frac{\partial\alpha^V}{\partial N}\bar{e}_i \geq \frac{\partial N\alpha^L\bar{e}_i}{\partial N}$.

(d) $0 \leq \frac{\partial\alpha^V}{\partial d} = \frac{c(2-\lambda)}{2d^2N(1-\lambda)\bar{e}_i} \leq \frac{2c}{2d^2N(1-\lambda)\bar{e}_i} = \frac{\partial\alpha^L}{\partial d}$. $\frac{\partial(N_P^V/N)}{\partial d} = \left(\alpha^V\right)^{-2}(\alpha^V\frac{\partial\alpha^L}{\partial d} - \alpha^L\frac{\partial\alpha^V}{\partial d}) \geq 0$. $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial d} = \frac{\partial N\alpha^L\bar{e}_i}{\partial d} + \frac{\partial\alpha^V}{\partial d}\bar{e}_i \geq \frac{\partial N\alpha^L\bar{e}_i}{\partial d}$.

(c) $0 \geq \frac{\partial\alpha^V}{\partial c} = \frac{-2+\lambda}{2dN^2(1-\lambda)\bar{e}_i} \geq \frac{2}{2dN^2(1-\lambda)\bar{e}_i} = \frac{\partial\alpha^L}{\partial c}$. $\frac{\partial(N_P^V/N)}{\partial c} = \left(\alpha^V\right)^{-2}(\alpha^V\frac{\partial\alpha^L}{\partial c} - \alpha^L\frac{\partial\alpha^V}{\partial c}) \leq 0$. $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial c} = \frac{\partial N\alpha^L\bar{e}_i}{\partial c} + \frac{\partial\alpha^V}{\partial c}\bar{e}_i \leq \frac{\partial N\alpha^L\bar{e}_i}{\partial c}$.

(λ) $0 \geq \frac{\partial\alpha^V}{\partial\lambda} = \frac{-c}{2dN(1-\lambda)\bar{e}_i} \geq \frac{-2c}{2dN^2(1-\lambda)\bar{e}_i} = \frac{\partial\alpha^L}{\partial\lambda}$. $\frac{\partial(N_P^V/N)}{\partial\lambda} = \left(\alpha^V\right)^{-2}(\alpha^V\frac{\partial\alpha^L}{\partial\lambda} - \alpha^L\frac{\partial\alpha^V}{\partial\lambda}) \leq 0$. $\frac{\partial N_P^V\alpha^V\bar{e}_i}{\partial\lambda} = \frac{\partial N\alpha^L\bar{e}_i}{\partial\lambda} + \frac{\partial\alpha^V}{\partial\lambda}\bar{e}_i \leq \frac{\partial N\alpha^L\bar{e}_i}{\partial\lambda}$. *QED*

Appendix B: Appendix for Chapter 3: Estimation results of WM IV and V models with $m = 3, 5,$ and 10

We show the estimation results of WM IV and V models with m=3, 5, and 10. We show the results of ISO 14001 adoption by water emitting facilities (Table B.1), water emission reduction (Table B.2), ISO 14001 adoption by air emitting facilities (Table B.3), and air emission reduction (Table B.4).

Table B.1: Estimation results of ISO 14001 adoption by water emitting facilities with different ’m’s

	Dependent variable:ISO 14001 adoption					
	WM IV			WM V		
	m=3	m=5	m=10	m=3	m=5	m=10
Workers/10^2	0.119***	0.096***	0.094***	0.134***	0.128***	0.109***
	(0.011)	(0.009)	(0.015)	(0.013)	(0.012)	(0.014)
Air emission intensity/10^6	0.056	0.031	0.075	0.018	0.034	0.029
	(0.270)	(0.265)	(0.271)	(0.270)	(0.263)	(0.254)
Water emission intensity/10^6	-0.254***	-0.145**	-0.137*	-0.208*	-0.144*	-0.147*
	(0.096)	(0.073)	(0.076)	(0.123)	(0.080)	(0.084)
Profit/revenue/10^3	1.11*	0.924	0.912	0.958	0.949	0.923
	(0.656)	(0.648)	(0.652)	(0.640)	(0.625)	(0.665)
Population density/10^3	-5.62	-5.08	-3.81	-4.83	-4.73	-4.36
	(3.47)	(3.51)	(3.57)	(3.52)	(3.41)	(3.56)
Income per capita/10^3	0.293	0.238	0.223	0.268	0.318	0.242
	(0.375)	(0.349)	(0.362)	(0.369)	(0.372)	(0.376)
Support	-0.038	-0.053	-0.032	-0.019	-0.022	-0.016
	(0.136)	(0.134)	(0.140)	(0.137)	(0.133)	(0.144)
Grievances against pollution	-0.500	-0.346	-0.612	-0.462	-0.558	-0.481
	(0.455)	(0.458)	(0.463)	(0.464)	(0.427)	(0.468)
Year dummy (2001)	-0.227***	-0.211***	-0.200***	-0.269***	-0.258***	-0.225***
	(0.075)	(0.074)	(0.073)	(0.082)	(0.085)	(0.078)
Constant	-19.8	-13.6	-29.6***	-32.7**	-34.2	-7.44
	(12.5)	(10.1)	(10.5)	(16.3)	(27.1)	(4.78)
ν	0.333***	0.357***	0.334***	0.048	0.043	0.166***
	(0.046)	(0.038)	(0.067)	(0.032)	(0.054)	(0.054)
Likelihood	-1237.74	-1411.22	-1121.92	-1101.20	-1465.35	-1099.74
Observation	1326	1326	1326	1326	1326	1326

All of models in this table are SAR probit models. *, **, and *** imply that the coefficient is significantly different from zero at the 1%, 5% and 1% levels, respectively.

Table B.2: Estimation results of water emissions reductions with different 'm's

	Dependent variable: Water emissions reductions (E^{W}_{t+1}/E^{W}_{t})					
	WM IV			WM V		
	m=3	m=5	m=10	m=3	m=5	m=10
Workers	-0.367	-0.368	-0.366	-0.367	-0.367	-0.367
	(1.73)	(1.73)	(1.73)	(1.73)	(1.73)	(1.74)
Profit/Revenue	-0.352	-0.352	-0.350	-0.351	-0.351	-0.355
	(4.51)	(4.51)	(4.52)	(4.52)	(4.52)	(4.52)
ISO	176	174	172	172	172	173
	(847)	(849)	(849)	(849)	(849)	(849)
Population density	-1171	-1167	-1167	-1167	-1167	-1169
	(1742)	(1745)	(1746)	(1746)	(1746)	(1746)
Income per capita	2.22	2.21	2.21	2.20	2.20	2.21
	(12.1)	(12.1)	(12.1)	(12.1)	(12.1)	(12.1)
Grievances against pollution	-1.92	-1.89	-1.90	-1.90	-1.90	-1.91
	(10.1)	(10.2)	(10.2)	(10.2)	(10.2)	(10.2)
Constant	-134	-134	-134	-135	-135	-135
	(487)	(488)	(488)	(488)	(488)	(489)
ρ	0.006	0.002	0.001	0.0003	0.0005	-0.003
	(0.000)	(0.000)	(0.000)	(0.0000)	(0.0000)	(0.000)
Log-likelihood	-6647.038	-6647.041	-6647.041	-6647.041	-6647.041	-6647.041
Observation	663	663	663	663	663	663

All of models in this table are fixed effects SAR models. *, **, and *** imply that the coefficient is significantly different from zero at the 1%, 5% and 1% levels, respectively.

Table B.3: Estimation results of ISO 14001 adoption by air emitting facilities with different 'm's

	Dependent variable: ISO 14001 adoption					
	WM IV			WM V		
	m=3	m=5	m=10	m=3	m=5	m=10
Workers/10^2	0.129***	0.107***	0.097***	0.122***	0.107***	0.085***
	(0.006)	(0.005)	(0.016)	(0.006)	(0.005)	(0.006)
Air emission intensity/10^6	0.167	0.16	0.163	0.169	0.154	0.165
	(0.129)	(0.263)	(0.130)	(0.127)	(0.130)	(0.125)
Water emission intensity/10^6	-0.127	-0.138	-0.139	-0.130	-0.121	-0.118
	(0.105)	(0.102)	(0.107)	(0.123)	(0.11)	(0.102)
Profit/revenue/10^3	0.003	-0.053	-0.050	0.097	0.072	0.061
	(0.248)	(0.220)	(0.220)	(0.236)	(0.229)	(0.233)
Population density/10^3	-1.78	-2.07***	-1.95	-1.65	-1.40	-1.20
	(1.149)	(1.184)	(1.24)	(1.18)	(1.11)	(1.17)
Income per capita/10^3	-0.150	-0.144	-0.216	-0.120	-0.113	-0.129
	(0.121)	(0.122)	(0.123)	(0.117)	(0.118)	(0.121)
Support	0.070	0.078	-0.054	0.068	0.065	0.059
	(0.050)	(0.049)	(0.046)	(0.049)	(0.046)	(0.048)
Grievances against pollution	-0.289*	-0.281	-0.235	-0.322	-0.345**	-0.355**
	(0.171)	(0.18)	(0.182)	(0.170)	(0.163)	(0.171)
Year dummy (2001)	-0.178 ***	-0.144***	-0.130 ***	-0.181***	-0.161	-0.131***
	(0.030)	(0.03)	(0.035)	(0.032)	(0.03)	(0.028)
Constant	-3.38 ***	-3.58***	-54.5 **	-4.03 **	-4.198***	-3.02 **
	(1.78)	(1.99)	(21.4)	(2.08)	(2.03)	(1.26)
ν	0.270***	0.419***	0.430***	0.199***	0.287***	0.417***
	(0.000)	(0.001)	(0.090)	(0.020)	(0.019)	(0.025)
Likelihood	-6658.74	-6614.50	-6352.29	-6523.54	-6461.20	-6337.96
Observation	7158	7158	7158	7158	7158	7158

All of models in this table are SAR probit models. *, **, and *** imply that the coefficient is significantly different from zero at the 1%, 5% and 1% levels, respectively.

Table B.4: Estimation results of air emissions reductions with different ’m’s

	Dependent variable: Air emissiions reductions (E^{A}_{t+1}/E^{A}_{t})					
	WM IV			WM V		
	m=3	m=5	m=10	m=3	m=5	m=10
Workers	0.0242	0.0249	0.0248	0.0248	0.0248	0.0256
	(0.0514)	(0.0515)	(0.0523)	(0.0528)	(0.0528)	(0.0526)
Profit/Revenue	0.004	0.0069	0.0062	0.0163	0.0162	0.0189
	(0.098)	(0.098)	(0.100)	(0.100)	(0.100)	(0.100)
ISO	-14.6	-13.8	-12.9	-13.1	-13.1	-13.3
	(20.6)	(20.1)	(20.4)	(20.6)	(20.6)	(20.5)
Population density	6.04	5.21	5.72	6.41	6.43	5.97
	(26.0)	(26.1)	(26.4)	(26.7)	(26.7)	(26.6)
Income per capita	-0.1705	-0.169	-0.1672	-0.178	-0.178	-0.1761
	(0.227)	(0.227)	(0.231)	(0.233)	(0.233)	(0.232)
Grievances against pollution	0.133	0.131	0.13	0.144	0.144	0.142
	(0.196)	(0.197)	(0.200)	(0.202)	(0.202)	(0.201)
Constant	6.44	6.06	6.03	6.72	6.73	6.36
	(10.5)	(10.5)	(10.7)	(10.8)	(10.8)	(10.7)
ρ	0.075^{***}	0.110^{***}	0.090^{***}	0.0001^{***}	-0.001^{***}	0.039^{***}
	(0.000)	(0.000)	(0.000)	(0.0000)	(0.000)	(0.000)
Likelihood	-25422.42	-25419.32	-25423.81	-25426.65	-25426.65	-25426.18
Observation	3579	3579	3579	3579	3579	3579

All of models in this table are fixed effects SAR models. *, **, and *** imply that the coefficient is significantly different from zero at the 1%, 5% and 1% levels, respectively.

Appendix C: Appendix for Chapter 4

C.1 Calculations of t^{HIC} and t^{LIC}

Suppose[1]

$$\frac{1}{2}F(\bar{e}_H) < C_H(e_H(c\bar{e}_L)) + c\bar{e}_L e_H(c\bar{e}_L) + \Omega_L(c\bar{e}_L). \tag{4}$$

t_L^{HIC} must satisfy

$$\frac{1}{2}F(\bar{e}_H) = C_H(e_H(t_L^{HIC})) + t_L e_H(t_L^{HIC}) + \Omega_L(t_L^{HIC}) \tag{5}$$

where $\frac{1}{2}F(\bar{e}_i) = t_i^C e_i^C(t_i^C) + C_i(e_i^C(t_i^C)) + \Omega_i^C(t_i^C)$. By rearranging (5),

$$((\alpha-1)c+\alpha)t^2 - 2c(\alpha\bar{e}_L - c\bar{e}_H)t + c^2\left\{\frac{\alpha}{1+c}\bar{e}_L^2 - F(\bar{e}_H)\right\} = 0 \tag{6}$$

Let $A_T = (\alpha-1)c+\alpha, B_T = c(\alpha\bar{e}_L - c\bar{e}_H)$, and $C_T = c^2\left[\frac{\alpha}{1+c}\bar{e}_L^2 - F(\bar{e}_H)\right]$. Then, $t_L^{HIC} = \frac{B_T+\sqrt{B_T^2-A_TC_T}}{A_T}$.

Next, suppose (4) does not hold. By rearranging (5),

$$t^2 - 2c\bar{e}_H t + c[F(\bar{e}_H) - 2\Omega_L(c\bar{e}_L)] = 0. \tag{7}$$

Let $C' = c[F(\bar{e}_H) - \Omega_L(c\bar{e}_L)]$. Then,

$t_L^{HIC} = c\bar{e}_H - \sqrt{(c\bar{e}_H)^2 - C'}$

because $0 \leq t_L^{HIC} \leq c\bar{e}_H$ and $C' > 0$. t_L^{HIC} exists as long as $(c\bar{e}_H)^2 - C' > 0$, i.e.,

---

[1] This is the condition for $t_L^{HIC} < c\bar{e}_L$.

$$\frac{1}{2}F(\bar{e}_H) < C_H(0) + \Omega_L(c\bar{e}_L). \tag{8}$$

Next, t_L^{LIC} must satisfy

$$C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H^C(t_H^C) = C_L(e_L(t_L^{LIC})) + t_L^{LIC} e_L(t_L^{LIC}) + \Omega_L(t_L^{LIC}) \tag{9}$$

if $C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H^C(t_H^C) < C_L(0) + \Omega_L(c\bar{e}_L)$. Otherwise, $t_L^{LIC} = c\bar{e}_H$.

C.2 Contribution under pooling equilibria when $(1+c)\bar{e}_L < E[\bar{e}]$ and when $t < c\bar{e}_L$ and $(1+c)\bar{e}_L \geq E[\bar{e}]$

In this case[2], the socially optimal pooling tax rate is the same as the socially optimal tax rate for the high-cost industry. Hence, the contribution should be

$$\Omega(t) = \begin{cases} \frac{\alpha}{2}\left[\frac{t}{c}^2 + p_L(\bar{e}_L - \frac{t}{c})^2 + (1-p_L)(\bar{e}_H - \frac{t}{c})^2 - \{p_L c\bar{e}_L^2 + (1-p_L)c/(1+c)\bar{e}_H^2\}\right] & if\ t < c\bar{e}_L \\ \frac{\alpha}{2}\left[p_L c\bar{e}_L^2 + (1-p_L)\{\frac{t}{c}^2 + (\bar{e}_H - \frac{t}{c})^2\} - \{p_L c\bar{e}_L^2 + (1-p_L)c/(1+c)\bar{e}_H^2\}\right] & otherwise. \end{cases} \tag{10}$$

If $(\alpha-1)c+\alpha \leq 0$, then the best pooling equilibrium tax rate is $t_P = 0$. Otherwise, the best pooling equilibrium tax rate is $t_P = t_H^*$ if $t_2^P > t_H^*$ (otherwise, $t_P = t_2^P$).

If $t > c\bar{e}_L$ and $(1+c)\bar{e}_L \geq E[\bar{e}]$, then the contribution should be

$$\Omega(t) = \frac{\alpha}{2}\left[p_L c\bar{e}_L^2 + (1-p_L)\{\frac{t}{c}^2 + (\bar{e}_H - \frac{t}{c})^2\} - \{p_L c/(1+c)\bar{e}_L^2 + (1-p_L)c/(1+c)\bar{e}_H^2\}\right]. \tag{11}$$

---

[2] This condition is the same as $c\bar{e}_L < \frac{(E[\bar{e}]\alpha - c\bar{e}_L)c}{(\alpha-1)c+\alpha}$

C.3 Contribution under pooling equilibria when $(1+c)\bar{e}_L < cE[\bar{e}]$ and when $q > \bar{e}_L$ and $(1+c)\bar{e}_L \geq cE[\bar{e}]$

If $(1+c)\bar{e}_L < cE[\bar{e}]$, the socially optimal pooling quota is the same as the socially optimal quota for the high-cost industry. Hence, the contribution should be

$$\Omega(q) = \frac{\alpha(1-p_L)}{2(1+c)}\left[(1+c)q - c\bar{e}_H^2\right]^2. \tag{12}$$

And, the best pooling equilibrium quota is q_2^P.

If $\bar{e}_L < q$ and $(1+c)\bar{e}_L \geq cE[\bar{e}]$, then the contribution should be

$$\Omega(q) = \frac{\alpha}{2}\left[p_L\bar{e}_L^2 + (1-p_L)(c(\bar{e}_H - q)^2 + q^2) - \{cE[\bar{e}^2] - \frac{(cE[\bar{e}])^2}{1+c}\}\right]. \tag{13}$$

C.4 Proof of Proposition 4.1

$$TC_q = C_i(q_i^C) + \Omega_i(q_i^C) = \frac{\alpha c}{2(1+c)A}\bar{e}_i^2. \tag{14}$$

$$TC_t = C_i(e_i(t_i^C)) + t_i^C e_i(t_i^C) + \Omega_i(t_i^C) = \begin{cases} \frac{\alpha}{2(1+c)}\bar{e}_i^2 & if\ \alpha \leq c \\ \frac{\alpha+2\alpha c - c - c^2}{2(1+c)A_T}\bar{e}_i^2 & otherwise \end{cases} \tag{15}$$

where $A = c + \alpha(1+c)$ and $A_T = -c + \alpha(1+c)$. If $\alpha \leq c$, $TC_q < TC_t$ because $c/A < 1$. Next, consider $\alpha > c$.

$$\begin{aligned} TC_t - TC_q &= \frac{\alpha + 2\alpha c - c - c^2}{2(1+c)A_T}\bar{e}_i^2 - \frac{\alpha c}{2(1+c)A}\bar{e}_i^2. \\ &= \frac{(\alpha + 2\alpha c - c - c^2)A - \alpha c A_T}{2(1+c)AA_T}\bar{e}_i^2 \\ &= \frac{(\alpha - c)(1+c)A + \alpha c^2 + (1+c)\alpha^2 c + \alpha c^2 - \alpha^2 c^2 - \alpha^2 c}{2(1+c)AA_T}\bar{e}_i^2 \\ &= \frac{(\alpha - c)(1+c)A + 2\alpha c^2}{2(1+c)AA_T}\bar{e}_i^2 > 0. \end{aligned}$$

Hence, $TC_t > TC_q$. *QED*

C.5 Proof of Proposition 4.2

proof of $\Leftarrow$

Suppose that a separating equilibrium exists but $t_H^C > \hat{t}$. Then, because there is a separating equilibrium, there is $\tilde{t}$ such that

$$C_H(e_H(t_H^C)) + t_H^C e_H(t_H^C) + \Omega_H(t_H^C) \leq C_H(e_H(\tilde{t})) + \tilde{t} e_H(\tilde{t}) + \Omega_L(\tilde{t}) \tag{16}$$

$$C_L(e_L(\tilde{t})) + \tilde{t} e_L(\tilde{t}) + \Omega_L(\tilde{t}) \leq C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H(t_H^C) \tag{17}$$

Because $t_H^C > \hat{t}$, $\Omega_L(\hat{t}) = \Omega_H(\hat{t}) > \Omega_H(t_H^C)$, and therefore, $\Omega_L(t) > \Omega_H(t_H^C)$ for all $t \geq \hat{t}$. if $\tilde{t} > t_H^C > \hat{t}$, then (17) does not hold because $\Omega_L(\tilde{t}) > \Omega_L(\hat{t}) > \Omega_H(t_H^C)$ and C_i and $te_i(t)$ are nondecreasing on T. Thus, $t_H^C > \tilde{t}$ must hold.

However, if $c\bar{e}_L \geq t_H^C$, we get $t_H^C \leq \tilde{t}$ by adding (17) to (16) because $C_L(e_L(t)) = C_H(e_H(t))$ and $te_L(t) - te_H(t) = t(\bar{e}_H - \bar{e}_L)$ when $t < c\bar{e}_L$. This is a contradiction.

If $\tilde{t} \leq c\bar{e}_L < t_H^C$, then by adding (17) to (16), we get

$$\begin{aligned}
\frac{(t_H^C)^2}{2c} + t_H^C(\bar{e}_H - \frac{t_H^C}{c}) - \frac{1}{2}c\bar{e}_L^2 &\leq \tilde{t}\Delta\bar{e} \\
t_H^C\Delta\bar{e} + \frac{1}{2}[\bar{e}_L(t_H^C - c\bar{e}_L) - \frac{t_H^C}{c}(t_H^C - c\bar{e}_L)] &\leq \tilde{t}\Delta\bar{e} \\
(t_H^C - \tilde{t})c\Delta\bar{e} &\leq \frac{1}{2}(t_H^C - c\bar{e}_L)^2.
\end{aligned}$$

Because $c\bar{e}_H > t_H^C$, $c\Delta\bar{e} > t_H^C - c\bar{e}_L$. Hence, $t_H^C - \tilde{t} < (t_H^C - c\bar{e}_L)/2$. By rearranging, $\tilde{t} > (t_H^C + c\bar{e}_L)/2 > c\bar{e}_L$. This is a contradiction.

If $c\bar{e}_L \leq \tilde{t} < t_H^C$, (16) implies $\Omega_L(\tilde{t}) = \Omega_L(c\bar{e}_L) \leq \Omega_H(t_H^C)$. $\Omega_L(\tilde{t}) \leq \Omega_H(t_H^C)$ and (17) imply that $t_H^C \leq \tilde{t}$. This is a contradiction. Therefore, $t_H^C \leq \hat{t}$ if there exists a separating equilibrium.

Proof of $\Rightarrow$

Because $t_H^C \leq \hat{t}$, there exists $\dot{t}$, $t_H^C \leq \dot{t} \leq \hat{t}$, such that $\Omega_H(t_H^C) \geq \Omega_L(\dot{t})$ and $C_H(e_H^C) + t_H^C e_H^C + \Omega_H(t_H^C) = C_H(\dot{e}_H) + \dot{t}\dot{e}_H + \Omega_L(\dot{t})$.

Case: $t_H^C \leq \hat{t} \leq c\bar{e}_L$

$$\begin{aligned}
C_L(e_L(\dot{t})) + \dot{t}e_L(\dot{t}) + \Omega_L(\dot{t}) &= C_H(e_H(\dot{t})) + \dot{t}e_H(\dot{t}) + \Omega_L(\dot{t}) - \dot{t}\Delta\bar{e} \\
&= C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H(t_H^C) - \dot{t}\Delta\bar{e} \\
&= C_L(e_L^C) + t_H^C e_L^C + \Omega_H(t_H^C) + (t_H^C - \dot{t})\Delta\bar{e} \\
&\leq C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H(t_H^C)
\end{aligned}$$

Hence, ICs for both types hold in this case.

Case: $\dot{t} \geq t_H^C \geq c\bar{e}_L$

$$\begin{aligned}
C_L(e_L(\dot{t})) + \dot{t}e_L(\dot{t}) + \Omega_L(\dot{t}) &= C_H(e_H(\dot{t})) + \dot{t}e_H(\dot{t}) + \Omega_L(\dot{t}) - \dot{t}e_H(\dot{t}) - C_H(e_H(\dot{t})) + C_L(e_L(t_H^C)) \\
&= C_H(e_H(t_H^C)) + t_H^C e_H(t_H^C) + \Omega_H(t_H^C) - \dot{t}e_H(\dot{t}) - C_H(e_H(\dot{t})) + C_L(e_L(t_H^C)) \\
&= C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H(t_H^C) \\
&\qquad - [C_H(e_H(\dot{t})) + \dot{t}e_H(\dot{t}) - C_H(e_H(t_H^C)) - t_H^C e_H(t_H^C)] \\
&\leq C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H(t_H^C)
\end{aligned}$$

Case: $\dot{t} \geq c\bar{e}_L \geq t_H^C$

$$\begin{aligned}
C_L(e_L(\dot{t})) + \dot{t}e_L(\dot{t}) + \Omega_L(\dot{t}) &= C_H(e_H(\dot{t})) + \dot{t}e_H(\dot{t}) + \Omega_L(\dot{t}) - \dot{t}e_H(\dot{t}) - \frac{1}{2c}(\dot{t}^2 - (c\bar{e}_L)^2) \\
&= C_H(e_H(t_H^C)) + t_H^C e_H(t_H^C) + \Omega_H(t_H^C) - \dot{t}e_H(\dot{t}) - \frac{1}{2c}[\dot{t}^2 - (c\bar{e}_L)^2] \\
&= C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H(t_H^C) - \frac{1}{2c}[\dot{t}^2 - (c\bar{e}_L)^2] - [\dot{t}e_H(\dot{t}) - t_H^C\Delta\bar{e}]
\end{aligned}$$

If $\frac{1}{2c}[\dot{t}^2 - (c\bar{e}_L)^2] + [\dot{t}e_H(\dot{t}) - t_H^C\Delta\bar{e}] \geq 0$,

$$C_L(e_L(\dot{t})) + \dot{t}e_L(\dot{t}) + \Omega_L(\dot{t}) \leq C_L(e_L(t_H^C)) + t_H^C e_L(t_H^C) + \Omega_H(t_H^C). \tag{18}$$

We show this.

$$\begin{aligned}\frac{1}{2c}[\dot{t}^2-(c\bar{e}_L)^2]+[\dot{t}e_H(\dot{t})-t_H^C\Delta\bar{e}] &= (\dot{t}-t_H^C)\Delta\bar{e}-\frac{1}{2}c\bar{e}_L^2-\frac{1}{2c}\dot{t}^2+\dot{t}\bar{e}_L \\ &= (\dot{t}-t_H^C)\Delta\bar{e}+\frac{1}{2}(\dot{t}-c\bar{e}_L)\bar{e}_L-\frac{1}{2c}\dot{t}(\dot{t}-c\bar{e}_L) \\ &= (\dot{t}-t_H^C)\Delta\bar{e}-\frac{1}{2c}(\dot{t}-c\bar{e}_L)^2\geq 0.\end{aligned}$$

This is because $\dot{t}-t_H^C\geq\dot{t}-c\bar{e}_L$, and $2c\Delta\bar{e}-\dot{t}+c\bar{e}_L=2c\bar{e}_H-\dot{t}-c\bar{e}_L\geq 0$. Hence, ICs for both types hold and there is a separating equilibrium. *QED*

C.6 Proof of Proposition 4.4

By rearranging (4.13) and (4.14),

$$f(q)=[c+(1+c)\alpha]q^2-2(c\bar{e}_L+\alpha c\bar{e}_H)q+c\bar{e}_L^2+\frac{\alpha c^2}{1+c}\bar{e}_H^2-G(\bar{e}_L)\geq 0 \tag{19}$$

$$g(q)=(1+c)\alpha q^2-2\alpha c\bar{e}_Hq+\frac{\alpha c^2}{1+c}\bar{e}_H^2-G(\bar{e}_L)\geq 0 \tag{20}$$

$$h(q)=[c+(1+c)\alpha]q^2-2(c\bar{e}_H+\alpha c\bar{e}_H)q+c\bar{e}_H^2+\frac{\alpha c^2}{1+c}\bar{e}_H^2-(G(\bar{e}_L)+H)\leq 0 \tag{21}$$

where $G(\bar{e}_i)=C_L(e_i(q_i^C))+\Omega_i^C(q_i^C)$ and $H=C_H(e_H(q_L^C))-C_L(e_L(q_L^C))$. We get (19) and (20) by rearranging (4.13) and get (21) from (4.14). When $q_H^C>\bar{e}_L$, the IC condition for the low-cost industry might be (20). Let $A=c+(1+c)\alpha, A''=(1+c)\alpha$, $B'=c\bar{e}_L+\alpha c\bar{e}_H$, $B=c\bar{e}_H+\alpha c\bar{e}_H$, $C'=c\bar{e}_L^2+\frac{\alpha c^2}{1+c}\bar{e}_H^2-G(\bar{e}_L)$, $C''=\frac{\alpha c^2}{1+c}\bar{e}_H^2-G(\bar{e}_L)$, and $C=c\bar{e}_H^2+\frac{\alpha c^2}{1+c}\bar{e}_H^2-(G(\bar{e}_L)+H)$. Then, a separaing equilibrium q must satisfy

$$q\geq q'=\frac{B'+\sqrt{B'^2-AC'}}{A}<\bar{e}_L \tag{22}$$

$$q\geq q''=\frac{\alpha c\bar{e}_H+\sqrt{(\alpha c\bar{e}_H)^2-A''C''}}{A''}=\frac{c}{1+c}\bar{e}_H+\frac{\sqrt{c/A}}{1+c}\bar{e}_L<\bar{e}_H \tag{23}$$

$$q\leq q_2=\frac{B+\sqrt{B^2-AC}}{A}. \tag{24}$$

where $f(q') = 0$, $g(q'') = 0$, and $h(q_2) = 0$. There exists q satisfying the IC condition for the high-cost industry because

$$\begin{aligned}
B^2 - AC &= A(G+H) + (c(1+\alpha)\bar{e}_H)^2 - (c+(1+c)\alpha)c\bar{e}_H^2 - \frac{(c+(1+c)\alpha)\alpha c^2}{(1+c)}\bar{e}_H^2 \\
&= A(G+H) - \frac{\alpha c}{1+c}\bar{e}_H^2 && (25)\\
&= (c+\alpha c)c\Delta\bar{e}^2 + \frac{\alpha c^2}{1+c}(\Delta\bar{e}^2 + 2\Delta\bar{e}\bar{e}_L) \\
&= (c+\alpha c)c\Delta\bar{e}^2 + \frac{\alpha c^2}{1+c}(\bar{e}_H^2 - \bar{e}_L^2) > 0 && (26)
\end{aligned}$$

where $G = \frac{\alpha c}{(1+c)A}\bar{e}_L^2$, $H = c\Delta\bar{e}^2 + 2\frac{\alpha c}{A}\Delta\bar{e}\bar{e}_L$, $\Delta\bar{e} = \bar{e}_H - \bar{e}_L$, and $\Delta\bar{e}^2 = (\bar{e}_H - \bar{e}_L)^2$. Suppose that q satisfying the IC condition for the high-cost industry is given by (20). Then, the smallest one is q'', $g(q'') = 0$, and $g(\bar{e}_L) \leq 0$. $h(q'') < 0$ because

$$g(\bar{e}_L) - h(\bar{e}_L) = -c\bar{e}_L^2 + 2c\bar{e}_H\bar{e}_L - c\bar{e}_H^2 + H = -c\Delta\bar{e}^2 + c\Delta\bar{e}^2 + 2\frac{\alpha c}{A}\Delta\bar{e}\bar{e}_L > 0 \quad (27)$$

$$\frac{\partial g}{\partial q} = 2(1+c)\alpha q - 2\alpha c\bar{e}_H \quad (28)$$

$$\frac{\partial h}{\partial q} = 2cq + 2(1+c)\alpha q - 2c\bar{e}_H - 2\alpha c\bar{e}_H = \frac{\partial g}{\partial q} + 2cq - 2c\bar{e}_H < \frac{\partial g}{\partial q} \quad \text{if } q < \bar{e}_H. \quad (29)$$

Hence, $q'' < q_2$.

Next, consider the case when the IC condition for the high-cost industry is given by (19). If $B'^2 - AC' < 0$, then any q satisfying (24) is a separating equilibrium. Next, we consider when $B'^2 - AC' \geq 0$.

$q' \leq q_2$ implies that there exists a separating equilibrium. Now, we show it.

$$B^2 - B'^2 = c^2\Delta\bar{e}(\bar{e}_H + \bar{e}_L + 2\alpha\bar{e}_H) \quad (30)$$

$$\begin{aligned}
AC - AC' &= A(c\Delta\bar{e}(\bar{e}_H + \bar{e}_L) - c\Delta\bar{e}(\bar{e}_H + \bar{e}_L - 2q_L^C)) && (31)\\
&= 2(Ac\Delta\bar{e}q_L^C) = 2c^2(1+\alpha)\Delta\bar{e}\bar{e}_L && (32)
\end{aligned}$$

$$B^2 - B'^2 - (AC - AC') = 2\alpha c^2\Delta\bar{e}(\bar{e}_H - \bar{e}_L) + c^2\Delta\bar{e}(\bar{e}_H - \bar{e}_L) = (2\alpha + 1)(c\Delta\bar{e})^2 > 0. \quad (33)$$

Hence, $q' \leq q_2$. *QED*