University of Alberta

Three Essays on Tradable Permits

by

Marian L. Weber

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**

Department of Economics

Edmonton, Alberta

Fall 1998

0-612-34858-X

University of Alberta

Library Release Form

Name of Author: Marian L. Weber

Title of Thesis: Three Essays on Tradable Permit Systems

Degree: Doctor of Philosophy

Year this Degree Granted: 1998

Marian L. Weber
8407-60 Street
Edmonton, Alberta
T6B 1M4

October 1, 1998.

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read an recommend to the Faculty of Graduate Studies and Research for acceptance a thesis entitled **Three Essays on Tradable Permits** submitted by **Marian L. Weber** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Robin Lindsey, Supervisor

Henry van Egteren

Wicktor Adamowicz

Grant Hauer

Brian Copeland
(External Examiner)

Sept. 28, 1998
Date

Abstract

Essay 1.

In this essay, we present a model of noncompliant firms operating in a marketable pollution permit market. In the main result of the paper, we show how compliance is related to the initial distribution of permits. Redistributing permits from a fixed total stock from a competitive fringe to a firm with market power will decrease cheating by the firm with market power and increase cheating by the competitive fringe, generating an ambiguous global impact on cheating.

Essay 2.

In this essay we develop a model of moral hazard between shareholders who wish to minimize abatement costs, and managers whose unobservable effort affects abatement costs. We find that there is an incentive for shareholders to invest in abatement capital to relax the incentive and participation constraints for controlling the manager. Total abatement costs are higher than under the first-best outcome. Simulations reveal that investment in abatement capital can be higher or lower than in the first-best case. The direction of change in investment is shown to be related to the elasticity of substitution between effort and investment in achieving low cost outcomes. Our model explains some anomalies between actual and predicted performance of the market for sulfur dioxide emissions legislated under the Clean Air Act Amendments (1990).

Essay 3.

In this essay the optimal allocation of water and pollutants along a river is developed as the solution to an optimal control problem. We investigate whether it is possible to implement the optimal solution through a decentralized decision making framework. We find that if users do not account for third party effects when bidding for permits (ie. users are myopic), then there is no policy that optimally allocates water and pollutants. However, if users are not myopic we can construct a Nash Equilibrium which is socially optimal when there are markets for both water rights and rights to pollute. The Nash Equilibrium supports location specific prices for both water licenses and pollution permits. Third party effects from water transfers are shown to be an equity rather than an efficiency issue.

To my parents

Acknowledgement

I would like to express sincere thanks to my committee, Dr. Robin Lindsey, Dr. Henry van Egteren, and Dr. Wiktor Adamowicz, and to the external committee members Dr. Grant Hauer, and Dr. Brian Copeland, for their effort on my behalf.

I would particularly like to thank Dr. Lindsey for his availability and guidance from which I have learned a great deal. I would also like to thank Dr. van Egteren for his help and for encouraging me to pursue the topic of my dissertation. Finally I would like to thank Dr. Adamowicz and Dr. Michael Percy for the support and advice they have provided throughout my graduate studies.

I would like to thank the Department of Economics and the Faculty of Graduate Studies and Research for financial assistance throughout my graduate program. I would also like to thank the support staff in the Department of Economics, Audrey Jackson, Charlene Hill, Margaret Howell, Leslie Wayne, and Louise Edwards for all of their help, and for maintaining ample supplies of licorice and chocolate.

Thanks to the many friends and colleagues who have made this journey memorable and rewarding.

Special thanks to Stephen Dobson for being there.

List of Tables

List of Figures

Introduction

Since Dales [4], the use of market-based instruments for regulating environmental externalities has become widely accepted in public policy. Tradable permits are currently used to regulate the allocation of air-borne pollutants, as well as water rights and water pollutants. Thirty years of investigation have contributed to our theoretical understanding of the static and dynamic efficiency properties of tradable permit systems. In particular, Montgomery [28] showed that when each firm in the permit market is required to hold a portfolio of licenses for pollution damages at a set of receptor points, a permit market minimizes pollution control costs independent of the initial allocation of permits. Krupnick et al. [17], and McGartland and Oates [26], clarified and expanded Montgomery's result to the case where firms hold licenses for emissions rather than damages. Downing and White [5], Malueg [25], Milliman and Prince [27], Jaffe and Stavins [13], and Jung et al. [14], explore the dynamic efficiency of permit markets, that is, whether they create the correct incentives for socially optimal investment and innovation. The evidence suggests that incentive based regulation outperforms technological standard setting, and that tradable permits in particular tend to do as well as, or outperform, other policy instruments in many settings. Nonetheless, the performance of tradable permit systems depends on their design, as well as the number of firms in the market and their behavior.

The aggregate cost of pollution control is sensitive to the initial design of permit systems. An important issue is the initial allocation of the permits. Auctioned permits (eg. [22,23]), may outperform grandfathered permit systems [10, 11, 27]. For example, if permits are grandfathered, the dynamic efficiency of the market depends on whether firms are net buyers or sellers [25]. Grandfathering also can generate market power [7, 11], although the effects appear to be of limited practical significance [32, 20], and may even be exploited to increase social welfare in a second-best environment [12]. Despite the problems which can arise from grandfathering the initial allocation, research based on the assumption of grandfathering is important since auctions are less politically attractive than grandfathering [8, 9, 20]. Finally, aggregate control costs can be sensitive to the initial distribution of permits due to transactions costs [31].

There are many examples in the literature which show that the rules governing transactions in the permits market can affect the performance of the trading system. Stand alone spot markets can create incentives for firms to over-invest in abatement capital because the government can not commit to future permit prices [18]. On the other hand, stand alone spot markets can reduce environmental innovation as the government is able to expropriate the rents from the innovation by offering permits as a competing instrument for abatement [19]. Investment decisions also can be inefficient when there is a positive probability of the regulator ratcheting the environmental standard as it learns about environmental damages [16]. Finally, the regulator may want to exclude some firms from the market when there is limited information [21].

The degree to which a firm is able to maximize profits will affect its compliance strategy, and therefore the regulator's ability to minimize pollution control costs. Tschirhart

[33] and Postner [29] suggest that hierarchical and information structures within firms may reduce a firm's ability to respond to market incentives and therefore increase pollution control costs. At the same time, the regulatory environment which determines cost-recovery for firm investment and compliance decisions can also lead to inefficient compliance strategies and increased control costs [1, 3].

Finally, the greatest obstacle to efficient regulation is limited information. First, the ability to meet an environmental standard is complicated by imperfect enforcement [15, 24]. Second, limited information restricts the applicability of tradable permit systems to point-source problems. That is, it must be possible to monitor, at least imperfectly, a firm's emissions and damages. However, many pollution problems result in damages to ambient quality which can not be attributed to specific sources. Due to information problems associated with non-point source pollution, the regulator must use instruments which induce the truthful revelation of firm emissions (see [30, 35]).

In this dissertation, we continue the above line of inquiry and further investigate problems in implementing marketable permit systems in the real world. Essay 1 concerns the question of the initial design of a permit market. We combine the frameworks of Hahn [7] and Malik [24] to analyse the effects of market power and non-compliance on emissions and violations in a tradable permit market. This essay is based on Van Egteren and Weber [34]. Due to an error in their model, which was pointed out by an anonymous author, Proposition 3 as well as the policy implications of Van Egteren and Weber [34] have been modified.[1] We show that net social welfare, defined as firm benefits less the sum of social damages and enforcement costs from emissions, is not invariant to the initial distribution of permits when there is both market power and non-compliance in the permit market. In the main propositions of the essay, we show that aggregate emissions and violations are a function of the initial endowment of permits to the firm with market power. Therefore the costs of enforcement depend on the initial endowment to the dominant firm. This suggests that the regulator can use permits as an implicit enforcement mechanism. Our model shows that the regulator needs to consider the interaction between market power, the initial endowment, and the enforcement parameters when designing welfare maximizing emissions trading systems.

As mentioned earlier, the structure of information and delegation of authority within a firm may reduce its ability to respond to market incentives. Moreover, it appears that while the emissions trading system initiated under Title IV of the Clean Air Act Amendments (1990) (CAAA) is a success, firms are overinvesting in abatement capital, and the total costs of compliance are not being minimized [2, 6]. In Essay 2, we use the theory of moral hazard to develop a theoretical basis for the inability of firms to respond optimally to the incentives of a permit market, and suggest that this may explain the overinvestment in abatement capital noted above. In our model, managerial effort and investment in abatement capital are both inputs in the compliance cost function. We show analytically that there can be a tendency to overinvest in abatement capital relative to the least cost

[1] We wish to thank the anonymous author for this contribution to the essay.

compliance strategy because moral hazard increases the cost of managerial effort. Further investigation of the model through numerical simulation reveals outcomes consistent with the stylized facts of the sulfur dioxide trading program under the CAAA.

In the last essay, we turn our attention from the use of tradable permits for regulating air pollution to the use of tradable permits for managing the allocation of effluent discharges and consumptive water rights along a river. In this paper, the efficiency of a market system for allocating water rights is complicated by directional externalities which are introduced when firms located along the river discharge pollutants. This is because water consumption and emission of pollutants upstream adversely impacts users downstream. This paper highlights the inefficiencies which result when multiple regulations with interdependent objectives are imposed on a system without consideration of the interdependence in the initial design. In this case, regulations for implementing water quality objectives can lead to unexpected outcomes when water is allocated by a market mechanism. In the main proposition of the paper, we show that a system of tradable permits for both consumptive water rights and pollution rights can lead to the efficient outcome. However, the information requirements and potential transactions costs required to generate the efficient outcome are significant and may be prohibitive.

In each of the three essays, we consider the performance of tradable permit systems when we relax the simplified assumptions of the textbook model. In the first essay we relax the assumption that the permit market is perfectly competitive. In the second essay we illustrate the problems associated with treating the firm as a "black box" when estimating pollution control costs. The model in the third essay illustrates the danger of ignoring competing management objectives when designing optimal regulations. These essays show that the outcome of a tradable permit system is sensitive to the assumptions of the model. This illustrates the continuing need to investigate the properties of tradable permit systems in real world settings.

References

[1] D. Bohi and D. Burtraw, Utility investment behavior and the emission trading market, *Resour. Energy* **14**, 129-153 (1992).

[2] D. Bohi and D. Burtraw, SO_2 allowance trading: how experience and expectations measure up, Resources for the Future, Discussion Paper 97-24 (1997).

[3] J. Coggins and V. Smith, Some welfare effects of emission allowance trading in a twice-regulated industry, *J. Environ. Econom. Management* **25**, 275-297 (1993).

[4] J. Dales, Land, water, and ownership, *Cdn. J. Econom.* **1**, 794-804 (1968).

[5] P. B. Downing and L. J. White, Innovation in pollution control, *J. Environ. Econom. Management* **13**, 18-29 (1986).

[6] A. D. Ellerman, R. Schmalensee, P.L. Joskow, J.P. Montero, and E.M. Bailey, Emissions trading under the U.S. Acid Rain Program: evaluation of compliance costs and allowance market performance, MIT, Center for Energy and Environmental Policy Research, Book #23 (1997).

[7] R. Hahn, Market power and transferable property rights, *Quart. J. Econom.* **99**, 735-765 (1984).

[8] R. Hahn, The political economy of environmental regulation: towards a unifying framework, *Public Choice* **65** 21-48 (1990a).

[9] R. Hahn, Regulatory constraints on environmental markets, *Journal of Public Economics* **42** 149-175 (1990b).

[10] R. Hahn and R. Noll, Designing a market for tradable emission permits, *in* "Reform of environmental regulation" (W. Magat, Ed.) Ballinger, Cambridge, MA (1982).

[11] R. Hahn and R. Noll, Barriers to implementing tradable air pollution permits: problems of regulatory interactions, *Yale J. Regul.* **1**, 63-91 (1983).

[12] R. Innes, C. Kling, and J. Rubin, Emission permits under monopoly, *Natur. Resour. Model.* **5**, 321-343 (1991).

[13] A. B. Jaffe, and R. N. Stavins, Dynamic incentives of environmental regulations: the effects of alternative policy instruments on technology diffusion, *J. Environ. Econom. Management* **29**, S43-S63 (1995).

[14] C. Jung, K. Krutilla and R. Boyd, Incentives for advanced pollution abatement technology at the industry level: an evaluation of policy alternatives, *J. Environ. Econom. Management* **30**, 95-111 (1996).

[15] A. Keeler, Noncompliant firms in TDP markets, *J. Environ. Econom. Management* **21**, 180-189 (1991).

[16] P. W. Kennedy, Learning about environmental damage: implications for emissions trading, Discussion Paper 97-2, Department of Economics, University of Victoria, Victoria, BC (1997).

[17] A.J. Krupnick, W.E. Oates, and E. VanDeBerg, On marketable air-pollution permits: the case for a system of pollution offsets, *J. Environ. Econom. Management* **10**, 233-247 (1983).

[18] J.-J. Laffont and J. Tirole, Pollution permits and compliance strategies, *Journal of Public Economics* **62**, 85-125 (1996a).

[19] J.-J. Laffont and J. Tirole, Pollution permits and environmental innovation, *Journal of Public Economics* **62**, 127-140 (1996b).

[20] J. O. Ledyard and K. Szakaly Moore, Designing organizations for trading pollution rights, *Journal of Economic Behavior and Organization* **25** 167-196 (1994).

[21] T. R. Lewis and D. E. M. Sappington, Using markets to allocate pollution permits and other scarce resource rights under limited information, *Journal of Public Economics* **57**, 431-455 (1995).

[22] R. Lyon, Auctions and alternative procedures for allocating pollution rights, *Land Economics* **58**, 16-32 (1982).

[23] R. Lyon, Equilibrium properties of auctions and alternative procedures for allocating transferable permits, *J. Environ. Econom. Management* **14**, 129-152 (1986).

[24] A. Malik, Markets for pollution control when firms are noncompliant, *J. Environ. Econom. Management* **18**, 97-106 (1990).

[25] D. A. Malueg, Welfare consequences of emissions credit trading programs, *J. Environ. Econom. Management* **18**, 66-77 (1990).

[26] A. M. McGartland and W. E. Oates, Marketable permits for the prevention of environmental deterioration, *J. Environ. Econom. Management* **12**, 207-228 (1995).

[27] S. R. Milliman and R. Prince, Firm incentives to promote technological change in pollution control, *J. Environ. Econom. Management* **17**, 247-265 (1989).

[28] W. Montgomery, Markets in licenses and efficient pollution control programs, *J. Econom. Theory* **5**, 395-418 (1972).

[29] H. H. Postner, Tradable-rights approach to environmental policy, Working Paper No. 30, Economic Council of Canada, Ottawa (1992).

[30] K. Segerson, Uncertainty and incentives for nonpoint pollution control, *J. Environ. Econom. Management* **15**, 87-98 (1988).

[31] R. N. Stavins, Transactions costs and tradable permits, *J. Environ. Econom. Management* **29** 133-148 (1995).

[32] T. Tietenberg, "Emissions trading: an exercise in reforming pollution policy", Resources for the Future, Washington D.C., (1985).

[33] J. T. Tschirhart, Transferable discharge permits and profit-maximizing behaviour, in "Economic Perspectives on Acid Deposition Control" (T.D. Crocker, Ed.), Butterworth, Boston, MA (1984).

[34] H. Van Egteren and M. Weber, Marketable permits, market power, and cheating, *J. Environ. Econom. Management* **30**, 161-173 (1996).

[35] A. P. Xepapadeas, Environmental policy under imperfect information: incentives and moral hazard, *J. Environ. Econom. Management* **20**, 113-126 (1991).

Essay 1

Marketable Permits, Market Power, and Cheating[1]

1. Introduction

Marketable pollution permits (e.g. [6]) are gaining acceptance in government policy circles. The theoretical arguments supporting these schemes are now widely known. Because all firms face the same price for a permit, emissions are adjusted in a manner which ensures that the marginal cost of abatement is equalized across firms. As a consequence, aggregate abatement costs are minimized for a given environmental standard. Furthermore, Montgomery [19] has shown that this result holds under any allocation rule used in the initial distribution of permits. Based on these results, initial estimates of potential cost savings for marketable permit systems over command and control regulations were deemed significant and achievable with minimal information requirements on the part of regulators (*c.f.*, [1-3], and [17]).

This compelling set of results is one reason why proponents of marketable pollution permits have been able to overcome the public stigma associated with selling rights to pollute. However, the original promise of marketable permit systems has not been fulfilled. Significant gains in pollution reduction have not been achieved, and emissions permits have become costly liabilities to firms (see [10]). In addition, the theoretical models supporting permits have been challenged. In order for the least cost result to hold, these models must

[1] A version of this chapter has been published. See [19].

maintain rigid assumptions about the inability of firms to cheat and about the level of market power of firms participating in the permits market (see [8, 12, 14]).

One implication of Montgomery's [16] work is that the initial allocation of permits is purely an equity issue. This is challenged by Hahn [8] who shows that the initial distribution of permits has efficiency implications when one firm perceives that it has the power to influence price in the permits market. In Hahn's model, if the initial endowment of a firm with market power is different from its equilibrium demand for permits, then aggregate abatement costs exceed aggregate abatement costs at the least cost solution. In addition, Hahn [8] demonstrates that the degree of distortion, or inefficiency, is increasing in the endowment of permits received by the firm with market power. This is important because it shows that the initial allocation itself affects the market power of firms participating in the permits market (see [9]). For instance, a firm with monopsony power understands that it faces a rising equilibrium price for permits as its demand for permits increases, and it takes this into consideration when determining how many permits to buy and how much pollution to abate. The initial allocation is important in determining how many permits a given firm purchases and therefore can either enhance or mitigate the perceived market power of the monopsonist. As a result, changing the initial endowment directly influences the efficiency properties of marketable permit systems.[2]

[2] It is important to note that the initial endowment will impact on market power only if the firm can credibly commit to a monopoly or monopsony pricing strategy. Since permits may be viewed as durable goods, we may have a durable goods monopoly pricing problem. If the monopoly firm cannot commit to the monopoly pricing strategy (a more thorough examination of the institutions is required here), then the perfectly competitive price will

Despite the *theoretical* demonstration that market power influences aggregate abatement costs, there is still debate over the *practical* importance of market power. Tietenberg [18] notes that in most studies involving simulation of abatement costs in permit markets with market power, aggregate abatement costs do not increase significantly even where there is aggressive manipulation of price. Moreover, market power per se should not affect environmental quality since, by design, permit markets hold the aggregate level of emissions constant. Thus market power does not seem to be a significant challenge to the practical importance of these systems.

The same can not be said for noncompliance. Noncompliance arises because limited budgets and prohibitive monitoring costs make complete enforcement impossible. Firms cheat if the marginal penalty for cheating is less than the marginal cost of compliance, which is simply the cost of obtaining a permit for an additional unit of emissions. Malik [14] shows that with noncompliance, transferable permits markets do not minimize abatement costs as changes in the permit price affect the number of cheating firms as well as the allocation of abatement costs across firms. Moreover, he argues that when the number of firms in the permit market is small, noncompliance on the part of one firm can have a significant effect on the equilibrium outcome through its effect on permit prices. This

---

result, regardless of the initial endowment.

Since we are dealing with a static model in which the market for permits disappears after one set of market clearing transactions, we do not have the monopolist selling at two different time periods. Thus, we avoid the problem by artificially inducing commitment. Nonetheless, we feel this is an important aspect of the problem needing further investigation. We thank Arthur De Vany for his keen insight in observing this aspect of the problem.

introduces a further distortion in the efficiency of the permit market unaccounted for in most of the literature.[3]

In this paper, we combine the theoretical challenges of both Malik [14] and Hahn [8] to consider the impact of market power on equilibrium permit prices and levels of compliance. This paper is a modified version of Van Egteren and Weber [19]. In particular, Proposition 3 has been modified to correct for an error identified by an anonymous author. The chief finding is that when a firm has market power in the permit market, the initial allocation is fundamental in determining prices and levels of compliance for all participants in the permit market. Furthermore, because monitoring and enforcement impose social costs, the exercise of market power and its corresponding impact on the equilibrium level of compliance is a significant factor in determining total social costs. In particular, we show that if there is a firm with market power (monopsonist buyer or monopoly seller, hereafter referred to as the dominant firm) in the permit market, then an increase in the endowment of permits to this firm, accompanied by an equivalent reduction in the endowment of the competitive fringe, leads to an unambiguous increase in global violations and global emissions as long as the dominant firm is compliant. When the dominant firm is noncompliant, then a redistribution of permits from the competitive fringe to the dominant firm generates an ambiguous impact on global violations and global emissions since a redistribution of permits from the competitive fringe to the dominant firm leads to an

[3] Fuller [7] shows that the role of enforcement in the choice of abatement strategies is empirically significant.

increase in the equilibrium price of permits. This increase in price leads to an increase in violations and a decrease in emissions among the fringe, but because the dominant firm holds more permits in equilibrium, its violations decrease even though its emissions increase. Since violations and emissions for the two groups move in opposite directions, the overall impact on global compliance and emission levels is parameter specific. Also, the change in net social benefits, comprised of firm profits, pollution damages, and enforcement costs, from increasing the endowment to the firm with market power is ambiguous.

The results generated by our model suggest that the initial allocation of permits might be used as a policy tool for increasing the performance of pollution permit markets. The initial allocation can be used not only to mitigate the possible exercise of market power by some firms, but also as an implicit enforcement mechanism. Decreasing the endowment of permits to the firm with market power will enhance the compliance level of the competitive fringe. However, this will generate an increase in the violations of a noncompliant firm with market power, suggesting that this is where monitoring efforts should be focused as it would seem less expensive to spend limited enforcement budgets on a single firm rather than on a myriad of firms in the competitive fringe.

The remainder of the paper proceeds as follows. In Section 3, we present a simple model of a noncompliant firm operating in a marketable pollution permits market. In Section 3, we analyze the optimal decisions of firms operating in a permits market in the presence of either monopoly or monopsony power. Policy implications are discussed in

Section 4, while Section 5 concludes the paper.

2. The Model.

Consider a market for tradable permits in which n firms participate. Each firm is endowed with some portion of a fixed stock of permits, $\bar{L}$. The fixed stock is determined by the environmental objective of the regulator. Firms are required to purchase permits, l_i, to cover their emissions, e_i. There is the possibility, however, that firms may cheat and emit more than their stock of permits allows, resulting in violations, $v_i = e_i - l_i$.

If firms cheat, then there is a firm-specific penalty function which consists of a probability that it will be audited and found in violation, $\beta_i(v_i)$, with $\beta_i' > 0$, $\beta_i'' \geq 0$, and $\beta_i(0) > 0$, and a fine whose magnitude depends upon the extent of the violation, $F_i(v_i)$, with $F_i' > 0$, $F_i'' \geq 0$, and $F_i(v_i) = 0$ for $v_i \leq 0$.[4] Thus, our analysis will consider only penalty schedules which have constant or increasing expected marginal penalties.[5]

A firm's final holdings of permits, l_i, depends in part upon the value of its output.

---

[4] Malik [14] shows that when the expected audit probability is a function of violations, it is possible to support the social optimum with noncompliance.

[5] For an analysis of fine schedules which are decreasing, see Keeler [12].

Following [14], we assume that the firm's optimal profit for a given level of emissions is

$$B_i(e_i) \equiv \max_{q_i} \; r_i q_i \; - \; C_i(q_i, e_i),$$

where r_i is the fixed price of firm i's output, q_i, and $C_i(\cdot,\cdot)$ captures costs of production. We assume that the benefit function is strictly concave in emissions so that $B_i' > 0$ and $B_i'' < 0$.

The firm with market power can be either a monopolist or a monopsonist. Throughout the paper, firm 1 will be the dominant firm. Firms 2 through I will be noncompliant while the final $K=n-I$ firms are assumed to be compliant. All firms are assumed to be risk neutral.

3. Firm Behavior and Market Equilibrium

3.1 The Competitive Fringe

A typical firm in the competitive fringe chooses emission levels and license holdings under conditions of perfect competition. A typical firm's choice problem is to

$$\text{(P0)} \qquad \underset{e_i, l_i}{maximize} \; B_i(e_i) \; - \; P \cdot (l_i - l_i^o) \; - \; \beta_i(v_i) F_i(v_i)$$

subject to $v_i \geq 0$, where P is the equilibrium permit price and l_i^o is firm i's endowment of permits. Buyers in the competitive fringe are distinguished from seller's by the sign of

$(l_i^o - l_i)$. This difference will be positive for sellers and negative for buyers.

If we assume that $e_i > 0$, which is the only interesting case, then the Kuhn-Tucker conditions for (P0) are:

(1) $$\frac{\partial \mathcal{L}}{\partial e_i} = B_i' - \beta_i' F_i - \beta_i F_i' + \lambda = 0,$$

(2) $$\frac{\partial \mathcal{L}}{\partial l_i} = -P + \beta_i' F_i + \beta_i F_i' - \lambda \leq 0,$$

(3) $$\frac{\partial \mathcal{L}}{\partial \lambda} = v_i \geq 0, \text{ and}$$

(4) $$l_i \geq 0;\ l_i \cdot \frac{\partial \mathcal{L}}{\partial l_i} = 0;\ \lambda \geq 0;\ \lambda \cdot \frac{\partial \mathcal{L}}{\partial \lambda} = 0,$$

where $\mathcal{L}$ is the Lagrange function formed from (P0) and λ is the Lagrange multiplier attached to the constraint.[6]

---

[6] Suppose now that condition (2) holds with equality. The second-order conditions require that the first principal minor of the Hessian determinant be negative, $|H_1| = B_i'' - \beta_i'' F_i - 2\beta_i' F_i' - \beta_i F_i'' < 0$; and that the Hessian determinant be positive,

3.2 The Compliance Decision.

Our analysis of the tradable permits market focuses on how the initial distribution of permits influences behaviour when firms may be noncompliant. When there is market power on the side of buyers or sellers, the initial allocation influences the equilibrium price of permits (see [8]). This is important because changes in the permit price will influence the number of firms that wish to comply with the environmental standard. That is, firms which were compliant before a price change may choose to be noncompliant after a price change, and vice versa.

In order to demonstrate how changes in the permit price impact on the compliance decision of firms in the competitive fringe, we will present a sufficient condition for noncompliance. This sufficient condition is the same as the one found in [14], except for some minor differences. For the firm to find it optimal to be compliant, condition (2) must hold as an equality when $v_i = 0$ since l_i would then be positive given that $e_i > 0$. Condition (2) becomes

(5) $$-P + \beta_i(0) F_i'(0) - \lambda = 0.$$

In order for this Kuhn-Tucker condition to be satisfied, $-P + \beta_i(0) F_i'(0)$ must be non-negative. Thus, $P \leq \beta_i(0) F_i'(0)$ specifies a necessary condition for compliance, while

---

$|H| = -B_i''(\beta_i'' F_i + 2\beta_i' F_i' + \beta_i F_i'') > 0$. Given the assumptions we have made, these requirements are satisfied.

a sufficient condition for noncompliance is to have $P > \beta_i(0)F_i'(0)$. A firm in the competitive fringe will be noncompliant if the equilibrium permit price is greater than the expected marginal penalty it would pay at the zero violation level. Given this condition, it is obvious that the compliance decision is independent of the initial allocation of a typical firm in the competitive fringe since each firm assumes it faces a fixed price. The only thing which will influence the compliance decision is a change in the equilibrium price, and this will depend upon the actions of the dominant firm.

The K compliant firms in the competitive fringe, each subscripted by k, will hold permits according to[7]

(6) $$l_k^* = B_k^{\prime -1}(P),$$

which emerges from the fact that $P = B_k'(l_k)$ when $v_k = 0$. This is obtained by adding conditions (1) and (2), when (2) holds with equality. Aggregating (6) over k then specifies the final holdings of permits by compliant firms. Note that, just as in Hahn [8], these holdings are independent of the initial distribution of permits. Therefore, for compliant firms, the only way final holdings of permits will change is if the price changes, and this again depends on the actions of the dominant firm.

For noncompliant firms, optimal choice of emissions and permit holdings is defined

[7] Throughout the paper, an asterisk denotes optimal choices.

implicitly by the first-order conditions resulting from (P0) with the constraint not binding, that is $P = B'(e_i^*) = \beta_i'(v_i^*)F(v_i^*) + \beta_i(v_i^*)F_i'(v_i^*)$.

Next, consider the following experiment. With the stock of permits fixed, take a small amount of the initial endowment of permits from a firm in the competitive fringe and give it to the firm with market power. This enables us to evaluate how the initial allocation influences the optimal violations for noncompliant firms.[8] In particular, we are looking for the sign of $\partial v_i^* / \partial l_1^o = (\partial v_i^* / \partial P) \cdot (\partial P / \partial l_1^o)$. The first part of this comparative static effect can be found from the solution to (PO). Given that

(7) $$\frac{\partial e_i^*}{\partial P} = - \frac{\beta_i'' F_i + 2\beta_i' F_i' + \beta_i F_i''}{|H|} < 0$$

and

(8) $$\frac{\partial l_i^*}{\partial P} = \frac{|H_1|}{|H|} < 0,$$

we have

(9) $$\frac{\partial v_i^*}{\partial P} = \frac{\partial e_i^*}{\partial P} - \frac{\partial l_i^*}{\partial P} = \frac{-B_i''}{|H|} > 0.$$

---

[8] Note that we do not consider the case of a redistribution of permits between firms in the competitive fringe as final demand is unaltered, so nothing changes.

Thus, because all firms in the competitive fringe solve the same optimization problem, the impact of changing the dominant firm's endowment influences the compliance levels of all firms. The direction of this influence depends upon the sign of $\partial P/\partial l_1^o$, which we explore next.

3.3 Market Equilibrium.

Market clearing requires that $\bar{L} = l_1^* + \sum_{i=2}^{l} l_i^* + \sum_{k=l+1}^{n} l_k^*$. If we let $\Delta \equiv \bar{L} - l_1$, then imposing market clearing implies that the equilibrium permit price, (the price which maximizes total profits for the dominant firm) is $P = p(\Delta^*)$, with $p' < 0$.[9] Hence,

$$(10) \qquad \frac{\partial P}{\partial l_1^o} = -p' \cdot \frac{\partial l_1^*}{\partial l_1^o}.$$

The sign of (10) depends upon how the optimal final demand for permits of the dominant firm is influenced by a change in the initial allocation of permits, $\partial l_1^*/\partial l_1^o$. In order to sign this partial derivative, we must determine the final demand for permits for the dominant firm.

---

[9] P is downward sloping because $B_i(e_i)$ is strictly concave. Linear demands, for example, would have $p'' = 0$.

3.4 The Monopoly/Monopsony Problem.

We assume that the dominant firm has sufficient market power to ensure that a monopoly seller (monopsony buyer) does not switch immediately to being a net buyer (resp. seller) when the price falls (resp. rises).[10] In addition, let $\delta \equiv l_1^o - l_1$, so that $\delta > 0$ (resp. $\delta < 0$) signals that the dominant firm is a monopolist (resp. monopsonist). Given these assumptions, the dominant firm attempts to

(P1) $$\underset{e_1, l_1}{maximize} \quad B_1(e_1) + p(\Delta)\cdot\delta - \beta_1(v_1)F_1(v_1)$$

subject to $v_1 \geq 0$. If we assume $e_1 > 0$, then the Kuhn-Tucker conditions for this program are

(11) $$\frac{\partial\mathcal{L}}{\partial e_1} = B_1' - \beta_1' F_1 - \beta_1 F_1' + \lambda = 0;$$

(12) $$\frac{\partial\mathcal{L}}{\partial l_1} = -p'\delta - p(\Delta) + \beta_1' F_1 + \beta_1 F_1' - \lambda \leq 0;$$

(13) $$\frac{\partial\mathcal{L}}{\partial \lambda} = v_1 \geq 0;$$

and

(14) $$l_1 \geq 0;\ l_1 \frac{\partial\mathcal{L}}{\partial l_1} = 0;\ \lambda \geq 0;\ \lambda \frac{\partial\mathcal{L}}{\partial \lambda} = 0.$$

For the firm with market power to find it optimal to be compliant, condition (12) must hold

[10] This ensures that the firm with market power has sufficient scope to manipulate price.

with equality when $v_1 = 0$, since l_1 would then be positive given $e_1 > 0$. Thus, a necessary condition for the firm with market power to be compliant is to have

$$-p'\delta - p(\Delta) + \beta_1(0)F_1'(0) \geq 0. \tag{15}$$

If the dominant firm is a monopolist, then (15) requires the marginal revenue from the sale of a permit to be less than or equal to the marginal expected penalty at the zero violation level.

If the dominant firm is compliant, so that (15) holds, then summing conditions (11) and (12), with (12) holding as an equality, yields

$$B_1' - p'\delta - p(\Delta) = 0. \tag{16}$$

Totally differentiating (16) and manipulating the expression gives

$$\frac{\partial l_1^*}{\partial l_1^o} = \frac{p'}{B_1'' + p''\delta + 2p'} > 0, \tag{17}$$

since $p' < 0$ and the second-order conditions require the denominator of this expression to be negative.[11] Equation (17) states that, as the dominant firm receives a larger endowment of permits, its final demand for permits increases. Hence,

Proposition 1: *If the firm with market power is compliant, then a redistribution of permits*

[11] Specifically, the second-order condition is to have $|\bar{H}| = -(B_1'' + p''\delta + 2p') > 0$.

from the competitive fringe to the firm with market power causes global violations to increase.

This happens because the noncompliant competitive fringe increases its violations as the equilibrium permit price increases. Recall that $\partial P/\partial l_1^o = -p' \cdot (\partial l_1^* / \partial l_1^o)$. Using (17) and the fact that $p' < 0$, we can establish that the equilibrium price rises as l_1^o rises for both the monopolist and the monopsonist. Combining this with equation (9) establishes that $\partial v_i^* / \partial l_1^o > 0$ for noncompliant firms.

The price increase also influences the behaviour of previously compliant firms in the competitive fringe. As the price rises, the "cost" of compliance also rises. With everything else constant, we see from (5) that this may cause some firms which were compliant to become noncompliant. Since the dominant firm remains compliant, an increase in violations by the competitive fringe means global violations have increased, thereby establishing the proposition.

In order to determine how the violations of a noncompliant firm with market power are influenced by a redistribution of the endowment, we analyze (P1) when the constraint is not binding. The first-order conditions for this problem are (11) and (12) without the

Lagrange multipliers.[12]

Given these requirements, and doing the necessary comparative statics establishes

Proposition 2: *If the dominant firm is noncompliant, then its own violations are decreasing in its endowment, while its emissions and licence holdings are increasing. In other words,* $\frac{\partial v_1^*}{\partial l_1^o} = \frac{-B''p'}{|H|} < 0$, $\frac{\partial e_1^*}{\partial l_1^o} = -\frac{p'(\beta_1'' F_1 + 2\beta_1' F_1' + \beta_1 F_1'')}{|H|} > 0$ *and*

$$\frac{\partial l_1^*}{\partial l_1^o} = -\frac{p'|H_1|}{|H|} > 0.$$

When the dominant firm is noncompliant, a redistribution of permits from the competitive fringe to the dominant firm generates an ambiguous impact on global violations and emissions. As the dominant firm receives more of the total endowment of permits, its market power changes. In the case of monopoly, the increased market power allows the firm to charge a relatively higher price to buyers. The higher price means that the marginal cost of compliance has increased for buyers and, all else equal, there will be an increased

---

[12] The second-order conditions require $|H_1| = B_1'' - \beta_1 F_1'' - 2\beta_1' F_1' - \beta_1 F_1'' < 0$, and $|H| = (\delta p'' + 2p')|H_1| - B_1''(\beta_1'' F_1 + 2\beta_1' F_1' + \beta_1 F_1'') > 0$. A sufficient condition for this to be true is to have $p'' \leq 0$, so that $\delta p'' + 2p' < 0$.

incidence of cheating among buyers. Furthermore, because the benefit function $B_i(e_i)$ is strictly concave in emissions, an increase in price leads to a decrease in the optimal level of emissions for both compliant and noncompliant buyers. The monopolist, however, holds more permits and therefore increases emissions while decreasing its overall level of violations. Since the impacts on violations and emissions are of opposing directions for the monopoly and the fringe, the net effect on global violations and emission depends on the relative shapes of the firm-specific benefit and fine functions.

In the case of monopsony, an increase in the monopsonist's endowment of permits leads to a reduction in its inherent market power. This produces a price for permits which is high relative to the price established before the transfer of endowments. Given that the monopsony buyer wants to purchase permits at as low a price as possible, the loss of market power inhibits its ability to achieve this objective. Once again, since prices are now higher, firms in the competitive fringe cheat more and emit less while the monopsonist cheats less and emits more so that the global impact on violations and emissions is ambiguous.

4. Policy Implications

The results generated by this simple model allow us to speculate on how market power might be exploited for policy purposes. With limited budgets and imperfect monitoring, the incentive to use the initial allocation of permits as an additional regulatory instrument may be strong. For example, because there is a firm which acts like a price setter in the market, adjusting the enforcement parameters facing this firm will affect aggregate levels of equilibrium emissions and compliance for all other firms participating in the market. This means that the regulator can achieve various outcomes simply by changing the initial allocation of the price-setting firm, rather than adjusting the policy parameters of all other firms.

Alternative allocation rules (see [13]) can be analyzed by comparing different endowments of permits to the dominant firm. Changing the endowment will influence equilibrium emissions and violations, thereby influencing net social benefits composed of firms' benefits, aggregate damages, and expected enforcement costs:

$$(18) \qquad NSB = \sum_{i=1}^{n} B_i(e_i^*) - D(e^*) - G\sum_{i=1}^{n} \beta_i(v_i^*).$$

We shall assume that damages, $D(e^*)$ are simply an increasing, convex function of aggregate emissions, where $\sum_{i=1}^{n} e_i^* = e^*$. Expected enforcement costs, $G\sum_{i=1}^{n} \beta_i(v_i)$, are composed of auditing costs, with each audit costing G dollars to perform.[13]

The regulator is interested in signing

$$(19) \qquad \frac{dNSB}{dl_1^0} = \sum_{i=1}^{n} B_i'(e_i^*)\cdot\frac{\partial e_i^*}{\partial l_1^0} - G\sum_{i=1}^{n} \beta_i'\cdot\frac{\partial v_i^*}{\partial l_1^0} - D'\cdot\frac{\partial e^*}{\partial l_1^0}.$$

[13] This net social benefit function reflects our assumption that fines are pure transfers. We thank a referee for making this clear to us.

Given that the competitive fringe sets $B_i'(e_i^*) = P$ for all i, we get

$$(20) \qquad \frac{dNSB}{dl_1^0} = [B_1'(e_1^*) - D'(e^*)]\frac{\partial e_1^*}{\partial l_1^0} + [P - D'(e^*)]\sum_{i=2}^{n}\frac{\partial e_i^*}{\partial l_1^0} - G\sum_{i=1}^{n}\beta_i' \cdot \frac{\partial v_i^*}{\partial l_1^0}.$$

Proposition 3: *If Firm 1 is a compliant monopolist, and* $B_1'(e_1^*) < D'(e^*) < P$, *then net social benefits are strictly decreasing in its endowment:* $\frac{dNSB}{dl_1^o} < 0$.[14]

From Proposition 1, $\partial v_i^*/\partial l_1^o > 0$ for each i, while $\partial e_i^*/\partial l_1^o = (\partial e_i^*/\partial P)\cdot(\partial P/dl_1^0) < 0$ from (7) and (10). If the dominant firm is compliant, then Eq. (17) establishes that $\partial e_1^*/\partial l_1^o > 0$, which, combined with the other assumption that $\beta_i'(v_i^*) > 0$, implies (20) is negative, and establishes the proposition.

In general the sign of (20) is ambiguous and depends on the price of permits, as well as the shape of the penalty and benefit functions. Proposition 3 shows that *given* a penalty structure and an initial distribution of permits, the regulator would be better off reducing the initial allocation to a compliant monopolist if equilibrium price exceeded marginal damages.

---

[14] We would like to thank an anonymous author for pointing out the correct approach to Proposition 3. Note that the condition in Proposition 3 can not hold for a monopsonist since $\delta < 0$ implies $B_1'(e_1^*) > P$.

Both the market structure and the level of violations in the permit market have a direct impact on net social benefits. Because enforcement costs can be large while budgets are limited, the regulator may wish to use the initial allocation of permits as a strategic policy instrument to reduce enforcement costs. Equation (20) shows the tradeoff between the initial allocation of permits to the dominant firm and enforcement costs. As the endowment to the compliant dominant firm increases, the level of violations of the competitive fringe, as well as total enforcement costs, increase. If this were the only effect then it would be optimal for the regulator to eliminate the market power of the compliant dominant firm. However, the increase in price accompanying an increase in l_1^0 reduces emissions of the competitive fringe, while increasing the emissions of the monopolist. If $B_1'(e_1^*) < D'(e^*) < P$, then a reduction in permits to the monopolist decreases the monopolist's emissions which is beneficial since $B_1'(e_1^*) < D'(e^*)$. At the same time, emissions of the fringe increase, which is also beneficial since $D'(e^*) < P$. Therefore society is strictly better off by decreasing the dominant firm's endowment.

In all other cases, the sign of (20) is ambiguous and depends on the relative size of the dominant firm with respect to the competitive fringe. To see this, note that we can rewrite (20) as

$$(21) \quad \frac{dNSB}{dl_1^0} = p'\delta \frac{\partial e_1^*}{\partial l_1^0} + [P - D'(e^*)]\left[\sum_{i=2}^{n} \frac{\partial e_i^*}{\partial l_1^0} + \frac{\partial e_1^*}{\partial l_1^0}\right] - G\sum_{i=1}^{n} \beta_i' \cdot \frac{\partial v_i^*}{\partial l_1^0}.$$

The first term in (21) is strictly negative for a monopolist, and strictly positive for a monopsonist. The sign of (21) depends on the sign of $P - D'(e^*)$, and whether the net change in emissions due to an increase in Firm 1's endowment is negative or positive. Finally, if the dominant firm is noncompliant, then it is not possible to sign the last term of (21) since its violations will be decreasing in its endowment. Ultimately, the interaction between these opposing effects determines the optimal distribution of permits to the dominant firm.

Ideally the regulator chooses l_1^0 so that (20) is equal to zero. However, because price is endogenous, the model is under-identified. The optimal choice of l_1^0 depends on the total number of permits in the market, the elasticity of demand for permits, and finally the penalty structure. A full model of the regulator's optimization problem is beyond the scope of this paper. As the discussion above indicates, the complex interactions between enforcement parameters and the initial allocation suggest that this is an area of analysis which needs to be explored further.

5. Conclusions

We have shown that the initial allocation of permits is an important tool for determining the performance of pollution permit markets when firms perceive that they have market power, and that the initial allocation can be used to vary the degree of market power in the permits market as in Hahn [8]. Moreover, we show that with noncompliance, market power distorts the equilibrium level of emissions and violations. Therefore, studies which ignore noncompliance when estimating the impact of market power in a permit market tend to underestimate the total social costs. This is contrary to the tidy results of earlier research suggesting that the properties of permit markets hold independent of the initial allocation. This independence property is attractive to governments wishing to implement permit markets since it minimizes the information requirement. In fact, this independence only holds under a set of ideal conditions. When the initial allocation matters, more information is required for implementing permits markets.

By analyzing the compliance decision of firms in a pollution permits market, we have shown that the initial allocation becomes crucial in determining permit prices, levels of compliance, levels of environmental quality, abatement costs and enforcement when firms perceive that they have market power. Our results suggest that the initial allocation should be viewed as a powerful policy instrument rather than an equity issue. In particular, the initial allocation complements other enforcement efforts. Since the impact of the initial allocation depends on the structure of the market as well as the structure of penalties, more research is required to understand the interdependence of these instruments. Finally, careful consideration of the structure of the permits market should be given before permits are allocated.

The main conclusion to be drawn from this analysis is that any deviation from perfect competition in a permits market with noncompliance can result in unexpected equilibrium outcomes. Without adequate information about the exact structure of the permits market, as well as knowledge of the social welfare function, the implementation of a permits market can lead unwittingly to a decline in social welfare relative to the ideal anticipated. Since these information requirements are non-trivial, an evaluation of the performance of permits markets against other regulatory instruments merits further research.

6. References

[1] R. Anderson, E. Seskin, and R. Reid, An empirical analysis of economic strategies for controlling air pollution, *J. Environ. Econom. Management* **10**, 101-121, (1983).

[2] S. Atkinson, Marketable pollution permits and acid rain externalities, *Cdn. J. Econom.* **16**, 704-722, (1983).

[3] S. Atkinson and T. Tietenberg, The empirical properties of two classes of designs for transferable discharge permit markets, *J. Environ. Econom. Management* **9**, 101-121, (1982).

[4] D. Bohi and D. Burtraw, Utility investment behavior and the emission trading market, *Resour. Energy* **14**, 129-153 (1992).

[5] J. Coggins and V. Smith, Some welfare effects of emission allowance trading in a twice-regulated industry, *J. Environ. Econom. Management* **25**, 275-297 (1993).

[6] J. Dales, Land, water, and ownership, *Cdn. J. Econom.* **1**, 794-804 (1968).

[7] D. Fuller, Compliance, avoidance, and evasion: emissions control under imperfect enforcement in steam-electric generation, *Rand J. Econom.* **18**, 124-137 (1987).

[8] R. Hahn, Market power and transferable property rights, *Quart. J. Econom.* **99**, 735-765 (1984).

[9] R. Hahn and R. Noll, Designing a market for tradable emission permits, *in* "Reform of environmental regulation" (W. Magat, Ed.) Ballinger, Cambridge, MA (1982).

[10] R. Hahn and R. Noll, Barriers to implementing tradeable air pollution permits: problems of regulatory interactions, *Yale J. Regul.* **1**, 63-91 (1983).

[11] R. Innes, C. Kling, and J. Rubin, Emission permits under monopoly, *Natur. Resour. Model.* **5**, 321-343 (1991).

[12] A. Keeler, Noncompliant firms in TDP markets, *J. Environ. Econom. Management* **21**, 180-189 (1991).

[13] R. Lyon, Equilibrium properties of auctions and alternative procedures for allocating transferable permits, *J. Environ. Econom. Management* **14**, 129-152 (1986).

[14] A. Malik, Markets for pollution control when firms are noncompliant, *J. Environ. Econom. Management* **18**, 97-106 (1990).

[15] W. Misiolek and H. Elder, Exclusionary manipulation of markets for pollution rights, *J. Environ. Econom. Management* **16**, 156-166 (1989).

[16] W. Montgomery, Markets in licenses and efficient pollution control programs, *J. Econom. Theory* **5**, 395-418 (1972).

[17] W. O'Neil *et al*, Transferable discharge permits and economic efficiency: the fox river, *J. Environ. Econom. Management* **10**, 346-55, (1983).

[18] T. Tietenberg, "Emissions trading: an exercise in reforming pollution policy", Resources for the Future, Washington D.C., (1985).

[19] H. Van Egteren, and M. Weber, Marketable permits, market power, and cheating, *J. Environ. Econom. Management* **30**, 161-173 (1996).

Essay 2

Moral Hazard and CAAA Compliance Strategies

1. Introduction

The 1990 Clean Air Act Amendments (CAAA) require electric utilities to reduce acid-rain emissions (sulfur dioxide and nitrogen oxide) by 50% from 1980 levels, a reduction of approximately 10 million tons per year. The policy is being implemented in two phases. Phase I, which took effect January 1995, requires that 263 of the highest emitting generating units meet an intermediate cap on emissions. Under Phase II, which takes effect January 1, 2000, all generating units will meet an aggregate 9-million ton per year cap on emissions. Under Title IV of the Act, the reduction of sulfur dioxide emissions is being achieved entirely through a national market in tradable allowances.

Initial evidence suggests that the emissions trading program is a success. Since January 1995, there has been significant over compliance. Permit prices and trading volumes have been lower than expected, while cost savings have been greater than expected (see [2, 7]). Actual average compliance costs for 1995 were \$187 per ton, while cost estimates prior to 1995 ranged from a low of \$180 per ton to a high of \$307 per ton [7].[1] Two main strategies are available for reducing emissions: investment in scrubbers (flue gas

---

[1] Part of the reason for the unexpected cost savings is railroad deregulation under the Staggers Act (1980) which significantly decreased the cost of supplying low sulfur coal from the Powder River Basin in Wyoming to generating units of the mid-west region where the majority of units affected by Phase I are located. The expansion of the geographic market for low sulfur coal into the Midwest has been accompanied by the introduction of cost-effective coal-blending power plants and falling costs of installing and operating scrubbers [2].

desulfurization), and fuel switching (switching to fuels with lower sulfur content including low sulfur coals, and/or natural gas and other petroleum products). Of these, scrubbing is the more capital intensive strategy.[2] About 50% of Phase I emission reductions are due to the installation of scrubbers. Bohi and Burtraw [2] argue that despite the unexpectedly low cost of compliance in 1995, potential cost savings from the market have not been achieved due to overinvestment in scrubbers relative to alternative compliance options.

In a static environment a marketable permit system provides correct incentives for socially optimal investment in abatement technology. Several models provide possible theoretical explanations for the bias towards scrubbers. Kennedy [12] shows that investment decisions can be distorted due to the inability of the regulator to commit to an aggregate level of emissions. However we don't believe that this argument is significant in the context of the CAAA since the emissions cap is legislated nationally and there would be significant political costs to changing the cap. Laffont and Tirole [14] argue that overinvestment may occur because raising public funds is costly. Therefore the regulator uses permit sales to raise revenue as well as to meet emissions objectives. In their model, stand alone spot markets induce excessive investment because the optimal price of pollution permits exceeds the marginal pollution cost in order to reduce the overall budget deficit. We also discount this reason for overinvestment. Under Phase I, the majority of permits are grandfathered to firms. Of the total allocation of permits, approximately 2.8% of total annual allowances are withheld each year to supply both the EPA spot and advance auctions [9].

---

[2] According to Bohi and Burtraw [2], the capital cost associated with switching from high to low sulfur coal is about 35% of the capital cost of a scrubber per kW^c of capacity.

Part of the bias towards scrubbing can be explained by regulatory distortions which favour investment in scrubbers over other compliance options. Electric utilities are regulated by public utility commissions. Therefore the compliance strategies of the utilities (and the potential cost savings from the allowance market) depend on the cost recovery rules applied by the respective public utility commissions, in particular the relative depreciation of compliance assets in the rate base (see [1, 5]). Yet Bohi and Burtraw [2] argue that, even when these regulatory distortions are accounted for, the bias towards scrubbers appears to be uneconomic *ex post*.

Electric utilities are complex hierarchical structures and attitudes towards risk are potential impediments to cost minimizing compliance strategies [22]. Observers have suggested that risk aversion on the part of managers who make compliance decisions may explain the preference of firms to build scrubbers rather than rely on permits [23]. In this paper we model the compliance strategy of a representative electric utility when there is moral hazard between shareholders of the firm (henceforth referred to just as the firm), and the manager who implements the compliance strategy. The moral hazard problem arises when risk averse managers can decrease abatement costs through the choice of some hidden but costly action which we will refer to as effort.

Managerial effort affects abatement costs in numerous ways. Managers expend effort in identifying and contracting sources of fossil fuels and allowances. They decide how to distribute intra firm emissions to minimize costs. For example, a manager can install a scrubber at one source and alter fuel inputs at another source, while shifting emissions to a third source. Finally the manager monitors the overall operations of the utility and presents the case for cost recovery at rate making hearings.

In order to encourage effort, the firm, in the model, makes the manager's wage utility contingent on the cost outcome. The extent to which the firm rewards or punishes the manager for good and bad cost outcomes is determined by two constraints. The participation constraint requires that the manager's expected utility over all outcomes be at least as high as some reservation level of utility. The extent to which the manager's utility is contingent on the state of the world is the power of the incentive. Since the manager is risk averse, while the firm is assumed to be risk neutral, it is optimal to insure the manager fully against negative outcomes. However, if effort is unobservable, then a fully insured manager supplies zero effort. The incentive compatibility constraint requires that the manager's choice of effort maximize her expected utility given the power of the incentive.

State contingent contracts are increasing in importance in the electric utilities industry. Since the energy cost shocks of the 1970s, public utility regulators have moved from cost plus regulation to incentive regulation which rewards (penalizes) firms with low (high) costs (see [18]). Evidence that managerial effort plays a large role in cost minimization under incentive regulation can be seen by looking at the structure of CEO salaries. As the regulatory environment has become more competitive, both the total magnitude and the percentage of incentive based components of managerial compensation have increased in the electric utility industry [4].[3]

The effects of moral hazard on the compliance strategies and compliance costs of firms are summarized as follows. It is well known that a first-best cost minimizing solution is not feasible with moral hazard [8]. In Proposition 1, we show that compliance costs under moral hazard are strictly greater than first-best-costs. We then determine whether moral hazard leads to a bias towards investment in scrubbers (referred to from now on as abatement capital) in the firm's compliance strategy. We show that under certain conditions, investment in abatement capital decreases the cost of managerial effort for the firm by relaxing both the participation and incentive compatibility constraints. In this sense, our results are consistent with those of other models where outside instruments are used to mitigate the effects of moral hazard (see [19, 3]).[4]

---

[3] Between 1990 and 1995, the magnitude of total CEO compensation for all industries increased by about 70%, while the level of incentive based compensation increased by more than 210% [4]. Still, the magnitude of CEO compensation, and its sensitivity to performance, are lower than average in the regulated utility industry. Recent studies suggest that this may be explained by the political constraints hypothesis, which maintains that public utility boards perform a monitoring role over managerial performance, and are under political pressure to keep CEO salaries low (see [10, 11, 4]). A possible alternative explanation for the lower salary profile is that there is a selection bias in the types of managers attracted to the electric utility industry; for example, CEOs are more likely to be engineers, and less likely to come from top schools in the regulated utility sector than the manufacturing sector [21].

[4] Note that while there are other ways to alleviate the moral hazard problem facing the firm, it can not be eliminated. For example, managerial performance at plants is not perfectly comparable because of heterogeneity between plants. Furthermore, hiring outside consultants and/or auditors is expensive. Therefore the firm still has an incentive to reduce the cost of the moral hazard constraint through input substitution.

In Proposition 2, we derive a condition in which investment decreases the magnitude of transfers to the manager in both high and low cost states of the world. This has the effect of reducing the cost of the participation constraint for a given level of effort. But this effect is weighed against the increased probability that the firm pays a higher transfer because investment increases the probability of a low cost outcome for the firm. The overall effect of investment on the participation constraint is ambiguous.

In Proposition 3, we show how the power of the incentive offered to the manager increases or decreases when there is an increase in investment according to whether investment and effort are substitutes or complements respectively in the production of low costs. If investment and effort are complements, an increase in investment leads to an increase in the marginal product of effort, and a decrease in the power of the incentive offered to the manager. There is a Pareto improvement in risk sharing between the firm and the manager, and the cost of the incentive constraint is reduced. However, if investment and effort are substitutes, investment leads to an increase in the power of the incentive, and a deterioration in risk sharing between the firm and the manager. In this case investment tends to increase the cost of the incentive constraint.

The overall impact of moral hazard on investment is ambiguous and depends on the technological relationship between investment and effort in the abatement cost function, as well as the shape of the manager's utility function. Our model is similar to the model of adverse selection presented in Laffont and Tirole [13], with effort replaced by a parameter reflecting hidden information. In their model, the regulator constructs a cost reimbursement rule for a firm undertaking an investment project. Investment costs depend on an efficiency parameter β which is revealed *ex post* to both the regulator and the firm. The firm has private information *ex ante* about a signal, θ, which is correlated with the efficiency parameter. The distribution of β is conditional on both θ, and the level of investment. As in our model, there is a correlation between the level of investment chosen by the firm and its *ex post* incentive scheme, but the linkage depends on the technological and informational parameters of the problem; in particular the interaction between θ and investment in the distribution function.

In general, investment choices are subject to both adverse selection and moral hazard. In this paper we ignore adverse selection and focus on the moral hazard aspect of the problem which exists solely because of risk aversion on the part of managers. This allows us to isolate the effect of moral hazard on the investment decision. The rest of the paper proceeds as follows. In section 2, we develop the formal model and graphically illustrate the incentive to over invest in abatement capital in order to reduce the costs of managerial effort. In section 3 we further investigate the effect using numerical simulations. Discussion and conclusions are presented in section 4.

2. Model

2.1 The Abatement Cost Function.

We define the abatement cost function as the difference in profits when emissions are constrained versus unconstrained (see [16]). Let the cost function $G(q, e)$ represent the minimum cost of producing output q and emissions e where $G_q > 0$ for all $q > 0$, and $G_e < 0$ for all $e < \bar{e}$, where $\bar{e}$ is the unconstrained optimal level of emissions. The cost function is convex in both e and q. Assume that the price of output p_q is fixed. Let $\bar{\pi} = p_q\bar{q} - G(\bar{q},\bar{e})$ be the maximum profit achievable when input decisions are unconstrained by environmental regulation. Let $\hat{\pi} = p_q\hat{q} - G(\hat{q},e)$ be the maximum profit which can be achieved given that emissions are constrained to be some level $e < \bar{e}$. The abatement cost function is just the difference in profits which results from constraining emissions,

$$C(e) = \bar{\pi} - \hat{\pi} = p_q(\bar{q} - \hat{q}) - G(\bar{q},\bar{e}) + G(\hat{q},e).$$

Differentiating the abatement cost function with respect to e yields

$$\frac{dC(e)}{de} = \left(\frac{\partial G}{\partial \hat{q}} - p_q\right)\frac{\partial \hat{q}}{\partial e} + \frac{\partial G}{\partial e}.$$

Since $\hat{q}$ is chosen optimally, $\frac{\partial G}{\partial \hat{q}} = p_q$. Therefore $\frac{dC(e)}{de} = \frac{\partial G}{\partial e}$ and $C(e)$ is convex by convexity of $G(.)$. Therefore, $C(e) > 0$, $C_e < 0$, and $C_{ee} > 0$. Assume that the cost function is stochastic and that there are two states of the world, either low cost, or high cost.[5] It is assumed that total abatement costs are lower and marginal abatement costs are lower in the low cost state of the world, so that $C^L(e) < C^H(e)$ and $C_e^L > C_e^H$.[6] Since $C^L(e) < C^H(e)$ for all e, then $\bar{e}_L < \bar{e}_H$.

---

[5] We have assumed a two-state model for analytical tractability. We expect the same qualitative results for a more general formulation with respect to differences in investment and total compliance costs between the first-best and second-best cases.

[6] Note that since $C_e < 0$, this implies that the benefit of an extra emission is lower in the low cost state of the world.

Total compliance costs for the firm are equal to abatement costs for reduced emissions plus the cost of holding permits for remaining emissions. The firm's objective is to minimize total compliance costs subject to the regulatory constraint, which in this case restricts the level of emissions to the number of permits held:

$$\min_{e,l} \; C^i(e) + p(l - l_o)$$

subject to

$$e - l = 0$$

where p is the price of permits, l is the number of permits which the firm holds in equilibrium, l_o is the firm's endowment of permits, and i denotes the state of the world. We assume that the permit market consists of a large number of firms with different technologies. Therefore it is assumed that the price of permits is independent of abatement costs for any one particular firm. [7] Substituting the constraint directly into the cost function, the first order condition for cost minimization is $p = -C_e$. We can think of $-C_e$ as the cost savings, or benefits, of an extra unit of emissions to the firm. This condition tells us that when $p < -C_e$, the price of a permit is less than the benefit of the extra unit of emissions and the firm will purchase a permit. Since the firm equates $-C_e^L(e_L^*) = p = -C_e^H(e_H^*)$ in both states of the world, $e_H^* > e_L^*$ by convexity of $C(e)$ and $C_e^L(e) > C_e^H(e)$.

Total compliance costs for the firm in each state of the world are given by:

---

[7] Note that because we have defined the permit market in terms of emissions rather than quality, the abatement cost function depends on the technology of the firm, and not on ambient air quality variables which would affect the clean-up costs of all firms in the market. The invariance of permit prices to the state of the world is more difficult to justify if permits are defined in terms of air quality.

$$c^L = pe_L^* + \int_{e_L^*}^{\bar{e}_L} -C_e^L(e)\,de; \text{ and}$$

$$c^H = pe_H^* + \int_{e_H^*}^{\bar{e}_H} -C_e^H(e)\,de.$$

Since $e_L^* < e_H^*$, $\bar{e}_L < \bar{e}_H$, and $C_e^L > C_e^H$, total compliance costs are lower for the firm in the low cost state of the world, $c^L < c^H$.

The probability of observing the low cost state of the world is assumed to be an increasing function of the level of investment in abatement technology, A, and the level of effort ϵ exerted by the manager. Assume that feasible levels of effort and investment are a compact subset of $\mathbb{R}^2$. The probability distribution function for costs is:

$$Pr(c = c^L) = \theta(\epsilon, A);$$

$$Pr(c = c^H) = 1 - \theta(\epsilon, A),$$

where $0 \leq \theta(\epsilon, A) \leq 1$.[8] The shape of $\theta(\epsilon, A)$ is governed by the following assumption:

Assumption 1 (A1): $\theta(\epsilon, A)$ *is strictly increasing and strictly concave in effort and investment.*

---

[8] Investment and effort enter the observed cost function indirectly as arguments in the probability distribution rather than directly as arguments of actual costs in each state of the world. This is a necessary simplification in the two-state model since making costs directly dependant on effort and investment would shift the support of the distribution, and the optimal level of effort could be guaranteed through a forcing contract. In the continuous model, the two formulations generate the same analytical results because the firm chooses effort and investment to minimize expected costs, and can't distinguish between effort and random shocks in observed costs.

We are employing the standard assumption that the firm becomes more efficient as extra units of investment and effort are added. The increased efficiency of the firm reduces the probability that a high cost outcome will occur. The marginal benefits of extra units of effort and investment are decreasing. This is captured by concavity of $\theta(\epsilon, A)$. Effort and investment can be either substitutes or complements in the probability distribution function. If $\theta_{\epsilon A} > 0$, then effort and investment are complements in the production of a low cost outcome. That is, an increase in investment leads to an increase in the marginal product of effort. If $\theta_{\epsilon A} < 0$, investment and effort are substitutes, that is, an increase in investment leads to a decrease in the marginal product of effort.

2.2 Managerial Incentive Problem

The manager's utility is assumed to be a positive function of income received t, and a negative function of effort. The manager is risk averse so that the marginal utility of income is strictly decreasing. We assume that the utility function is additively separable in income and effort. This implies that the choice of effort will be independent of the marginal utility of income of the manager, and vice-versa. These properties are summarized in Assumption 2:

Assumption 2 (A2): *The manager has a von Neumann-Morgenstern utility function* $U(t, \epsilon) = u(t) - v(\epsilon)$ *where* t *is income transferred to the manager,* $u(t)$ *is the wage-utility level of the manager, and* $v(\epsilon)$, *is the disutility of effort, where* $\epsilon \geq 0$. *We assume that* $v(\epsilon)$ *is increasing and strictly convex while* $u(t)$ *is continuous, increasing and strictly concave* *with domain **R**.*

Given (A2), we can define the inverse function $h \equiv u^{-1}$ which is strictly convex since $u(t)$ is strictly concave. The manager's reservation utility is equal to $\bar{U}$. The reservation price for a particular level of effort is given by $t^R(\epsilon) = h(\bar{U} + v(\epsilon))$.

The objective of the firm is to minimize expected total abatement costs. It is assumed that the firm can not directly observe the level of effort of the manager. However, total abatement costs are a signal of the manager's effort. Therefore, given the signal c^L or c^H, the manager is offered an incentive contract which specifies payments in either state of the world, $t_L = t(c^L)$, or $t_H = t(c^H)$, and generates corresponding wage-utility levels $u_L = u(t_L)$ and $u_H = u(t_H)$. We assume that investment is *ex ante* contractible. Therefore the firm can offer the manager an enforceable contract to choose a given level of investment prior to the state of nature being revealed. This allows us to examine whether or not the investment decision of the firm is distorted by the moral hazard problem.

The sequence of actions is as follows. The firm offers a contract $t(c)$ to the manager specifying payments to the manager based on the actual observation of abatement costs. Since investment is contractible, we assume that the firm chooses the level of investment at the same time the contract is offered to the manager. The manager then chooses a level of effort ϵ. Abatement costs are realised and permits are traded on the market.[9]

2.3 Benchmark Solution

In this section we derive the solution to the first-best case where effort is observable. We then use the first-best solution as a benchmark by which to compare the solution to the second-best case.

Since effort is observable in the first-best case, the firm can specify and enforce the manager's level of effort. The problem for the principal is to choose ϵ, A and t to minimize expected costs subject to the participation constraint of the manager:

(P1) $$\min_{\epsilon, A, t(.)} E\,(c + t(c) + rA)$$

subject to

(P1.1) $$E\,(u(t(c)) - v(\epsilon)) \geq \bar{U}; \qquad (\lambda \geq 0).$$

The Lagrangian for the problem is

$$\mathcal{L} = \theta(\epsilon, A)(c^L + t(c^L)) + (1 - \theta(\epsilon, A))(c^H + t(c^H)) + rA$$
$$- \lambda[\theta(\epsilon, A)\,u(t(c^L)) + (1 - \theta(\epsilon, A))u(t(c^H)) - v(\epsilon) - \bar{U}].$$

The first order conditions are given by

---

[9] In the model, the choice of emissions and the number of permits is independent of the level of effort and investment. That is, once the state of the world has been realised, the marginal abatement cost function is determined and the choice of emissions and permit holdings is given by the price of permits on the market. In this way, the equilibrium level of emissions and permits reveals the state of the world to the firm.

(1) $\mathcal{L}_\epsilon = \theta_\epsilon(\Delta c + \Delta t) - \lambda[\theta_\epsilon \Delta u - v'(\epsilon)] = 0;$

(2) $\mathcal{L}_A = \theta_A(\Delta c + \Delta t) + r - \lambda[\theta_A \Delta u] = 0;$

(3) $\mathcal{L}_{t_L} = 1 - \lambda u'(t_L) = 0;$

(4) $\mathcal{L}_{t_H} = 1 - \lambda u'(t_H) = 0;$

(5) $\mathcal{L}_\lambda = \theta(\epsilon, A)\, u(t_L) + (1 - \theta(\epsilon, A))u(t_H) - v(\epsilon) \geq \bar{U}; \quad \mathcal{L}_\lambda \lambda = 0;$

where

$\Delta c = c^L - c^H < 0;$

$\Delta t = t_L - t_H;$ and

$\Delta u = u(t_L) - u(t_H).$

Let ϵ^*, A^*, t_L^*, and t_H^*, represent the first-best solution. Conditions (3) and (4) imply that the marginal utility of income is equalized in both states of the world. Therefore the manager is offered a constant wage, $t^* = t_L^* = t_H^*$, satisfying the participation constraint such that

(6) $t^*(\epsilon^*) = h(\bar{U} + v(\epsilon^*)).$

(7) $-\theta_{\epsilon^*} \cdot \Delta c = \dfrac{v'(\epsilon^*)}{u'(t^*(\epsilon^*))};$ and

(8) $-\theta_{A^*} \cdot \Delta c = r.$

Solving (7) and (8) simultaneously determines the optimal level of effort ϵ^* and investment A^* for the principal. We can summarize the results under the benchmark solution. Since $t_L^* = t_H^*$, managers are fully insured. Since $\lambda > 0$, the agent receives no rent and $t^*(\epsilon^*) = t^R(\epsilon^*)$.[10] The optimal choice of effort equates the expected benefit from increased effort to the marginal disutility of the agent from increased effort (weighted by the marginal utility of income). Finally, the firm chooses the socially optimal level of investment by equating the expected marginal decrease in costs to the marginal cost of capital.

2.4 Unobservable Effort

If effort is unobservable then the manager chooses a level of effort which maximizes expected utility. The firm's problem with unobservable effort (P2) is analogous to (P1) with the addition of the incentive constraint (P2.2):

(P2) $\underset{\epsilon, A, t(.)}{Min} \quad E(c + t(c) + rA)$

subject to

(P2.1) $E(u(t(c)) - v(e)) \geq \bar{U};$ $\qquad (\lambda \geq 0);$

(P2.2) $\epsilon = argmax \ \theta(\epsilon, A)u(t_L) + (1 - \theta(\epsilon, A))u(t_H) - v(\epsilon)$ $\qquad (\mu \geq 0).$

Lemma 1 *(A3)*: *There exists a cost minimizing solution to (P2).*

This Lemma is formally proven in Proposition 1, Grossman and Hart [8].

If the expected utility of the manager is not concave in effort, then the level of effort which solves (P2.2) is not unique in general. This may occur, despite the concavity of the probability distribution function, if the optimal transfer function is not monotonic. Therefore

---

[10] Note that if $\lambda = 0$, the participation constraint is not binding. Therefore the only interesting case is when $\lambda > 0$.

the first order condition to (P2.2) may not even be a necessary condition for the cost minimization problem facing the firm. Lemma 2 proves that, given the assumptions of our model, (P2.2) can be replaced by the first order condition

(P2.2') $\theta_\epsilon[u(t_L) - u(t_H)] - v'(\epsilon) = 0;\ (\mu \geq 0)$.

We then show that at the optimum, both constraints are binding. Proofs of these results are established in the literature, but are included here for completeness.[11] In Proposition 1 we show that second-best abatement costs are strictly greater than first-best abatement costs.

Lemma 2: *If (A1) and (A2) hold, then (P2.2) can be replaced with (P2.2').*

Proof: First we show that the transfer function is non-increasing in costs. Suppose to the contrary that $t_H > t_L$. Then expected utility is maximized by setting effort equal to zero. But if effort is equal to zero then, given the risk aversion of the manager, the firm can do better by fixing the transfer in each state at the reservation utility so that $t_H - t_L = 0$, a contradiction. Hence $\Delta t \geq 0$, and $\Delta u \geq 0$.

Now $$\frac{\partial E\ u(t(c))}{\partial \epsilon} = \theta_\epsilon \Delta u - v'(\epsilon),$$

and $$\frac{\partial^2 E\ u(t(c))}{\partial \epsilon^2} = \theta_{\epsilon\epsilon} \Delta u - v''(\epsilon)$$

which is strictly negative by Assumptions (A1) and (A2). Therefore, $E(u(t))$ is concave in effort and the first order condition is necessary and sufficient for a unique global optimum

[11] See Grossman and Hart [8].

for the manager's level of effort. This implies that we can replace (P2.2) with (P2.2'). *Q.E.D.*

We can now characterize the solution for (P2). The Lagrangian for (P2) is given by

$$\mathcal{L} = \theta(\epsilon,A)(c^L + t(c^L)) + (1-\theta(\epsilon,A))(c^H + t(c^H)) + rA$$
$$-\lambda[\theta(\epsilon,A)u(t(c^L)) + (1-\theta(\epsilon,A))u(t(c^H)) - v(\epsilon) - \bar{U}].$$
$$-\mu[\theta_\epsilon(u(t(c^L)) - u(t(c^H))) - v'(\epsilon)].$$

The first order conditions for (P2) are:

(9) $\mathcal{L}_\epsilon = \theta_\epsilon(\Delta c + \Delta t) - \lambda[\theta_\epsilon \Delta u - v'(\epsilon)] - \mu[\theta_{\epsilon\epsilon}\Delta u - v''(\epsilon)] = 0;$

(10) $\mathcal{L}_A = \theta_A(\Delta c + \Delta t) + r - \lambda\theta_A \Delta u - \mu\theta_{\epsilon A}\Delta u = 0;$

(11) $\mathcal{L}_{t_L} = \theta(\epsilon,A) - \lambda\theta(\epsilon,A)u'(t_L) - \mu\theta_\epsilon u'(t_L) = 0;$

(12) $\mathcal{L}_{t_H} = (1-\theta(\epsilon,A)) - \lambda(1-\theta(\epsilon,A))u'(t_H) + \mu\theta_\epsilon u'(t_H) = 0;$

(13) $\mathcal{L}_\lambda = \theta(\epsilon,A)u(t_L) + (1-\theta(\epsilon,A))u(t_H) - v(\epsilon) \geq \bar{U}; \quad \mathcal{L}_\lambda \lambda = 0;$

(14) $\mathcal{L}_\mu = \theta_\epsilon \Delta u - v'(\epsilon) = 0.$

Let $\tilde{\epsilon}$ and $\tilde{A}$ be the second-best choice of effort and investment. In Lemma 3 we show that the participation constraint is binding. As a corollary which will be useful later on, we rank transfers in the low and high cost state of the world relative to the first-best case. Since both the participation and the incentive constraint are binding, expected abatement costs are higher relative to the benchmark case. This result is established in Proposition 1.

Lemma 3: *The participation constraint (P2.1) is binding.*

Proof: Consider, to the contrary, an incentive $u' = (u_L', u_H')$ which satisfies the constraints of (P2) where $\theta(\epsilon, A) u_L' + (1 - \theta(\epsilon, A)) u_H' - v(\epsilon) > \bar{U}$. Then there exists an incentive $u'' = (u_L' - \alpha, u_H' - \alpha)$ where $\alpha > 0$ is small, which still satisfies (P2.1) and (P2.2'). Therefore $u' = (u_L', u_H')$ does not minimize costs. *Q.E.D.*

Corollary: *For a given level of effort and investment,* $(\tilde{\epsilon}, \tilde{A})$, *the payment to and utility level of the manager in high and low cost states of the world can be ranked as follows:* $t_L > h(\bar{U} + v(\tilde{\epsilon})) > t_H$, *and* $u_L > u(h(\bar{U} + v(\tilde{\epsilon}))) > u_H$.

Proof: By Lemma 2 either $u_L > u_H$, or $\tilde{\epsilon} = 0$. We rule out the possibility that $\tilde{\epsilon} = 0$, since this case is uninteresting. Lemma 3 implies $\theta(\epsilon, A) u_L + (1 - \theta(\epsilon, A)) u_H = \bar{U} + v(\tilde{\epsilon})$ which together with *A2* proves the corollary. *Q.E.D.*

Proposition 1: *Second-best costs to the firm are strictly greater than first-best costs for all* $\epsilon^* \neq 0$:

Proof: First-best and second-best total abatement costs are defined as $TC_{FB}(\epsilon, A) = h(\bar{U} + v(\epsilon)) + rA$, and $TC_{SB}(\epsilon, A) = \theta(\epsilon, A) h(u_L) + (1 - \theta(\epsilon, A)) h(u_H) + rA$, respectively. From Lemma 3, the participation constraint in the second-best problem is binding therefore $\bar{U} + v(\epsilon) = \theta(\epsilon, A) u_L + (1 - \theta(\epsilon, A)) u_H$. Now $u_L > u_H$, since otherwise $\epsilon = 0$, which we have ruled out. Convexity of h then implies $\theta(\epsilon, A) h(u_L) + (1 - \theta(\epsilon, A)) h(u_H) > h(\bar{U} + v(\epsilon))$, while the cost of implementing A remains rA. This proves that $TC_{SB}(\tilde{\epsilon}, \tilde{A}) > TC_{FB}(\tilde{\epsilon}, \tilde{A}) \geq TC_{FB}(\epsilon^*, A^*)$. *Q.E.D.*

We are now in a position to examine the effect of moral hazard on investment in pollution abatement technology, in particular under what conditions we can expect investment in abatement technology to increase relative to the first-best outcome. The first order conditions for (P2) simultaneously determine $(\tilde{\epsilon},\tilde{A})$. In general $\tilde{\epsilon} \neq \epsilon^*$ and $\tilde{A} \neq A^*$. We can explain the change in A from the first-best to the second-best case from the perspective of production theory. There are two opposing effects. On the one hand, the incentive constraint makes effort more costly which encourages the firm to substitute away from effort. On the other hand, the marginal cost of producing output also increases. This encourages the firm to reduce output, which leads to a reduction in both inputs. Both effects lead to a decrease in ϵ, while A increases or decreases depending on whether the substitution effect outweighs the output effect. Production theory, however, does not tell the whole story. In production theory, the cost of inputs is exogenous. With moral hazard, the cost of effort is endogenous, and depends on A. Therefore, the final change in A depends on the interaction between investment and the cost of implementing a given level of effort. We examine this effect in the remainder of this section.

We can analyse the interaction between investment and effort by determining whether $\tilde{A} > A^*$ for a given value of effort, ϵ^*. That is we compare (10) to (8) evaluated at ϵ^*:

$$-\theta_A(\epsilon^*,A^*)\Delta c - r = 0; \text{ and,}$$

$$-\theta_A(\epsilon^*,\tilde{A})(\Delta c + \Delta t) - r + \lambda\theta_A(\epsilon^*,\tilde{A})\Delta u + \mu\,\theta_{\epsilon A}(\epsilon^*,\tilde{A})\Delta u = 0.$$

These conditions give the marginal expected profit from increasing A in the first-best and second-best cases respectively holding effort constant. By examining these two conditions, we see that A increases or decreases depending on the sign of

(15) $\theta_A(\epsilon^*, A^*)\Delta t - \lambda\theta_A(\epsilon^*, A^*)\Delta u - \mu\,\theta_{\epsilon A}(\epsilon^*, A^*)\Delta u.$

Expression (15) reflects the change in the cost to the manager of implementing the level of effort ϵ^*. If (15) is negative (resp. positive), then from the second order condition $\partial^2\mathcal{L}/\partial A^2 > 0$, we require $\tilde{A}(\epsilon^*) > A^*(\epsilon^*)$, (resp. $\tilde{A}(\epsilon^*) < A^*(\epsilon^*)$). In general the sign of (15) is indeterminate, and the change in investment depends on the parameters of the problem. In the next two propositions we will show that the firm can use investment strategically to reduce the cost of implementing a particular level of effort through the participation and incentive constraints. The costs of these constraints to the firm are reflected in the multipliers λ, and μ.

The magnitude of utility in the low and high cost states determines whether or not the participation constraint is satisfied. From (13) and (14),

(16) $u_L = \bar{U} + v(\epsilon) + (1-\theta)v'(\epsilon)/\theta_\epsilon;$

(17) $u_H = \bar{U} + v(\epsilon) - \theta v'(\epsilon)/\theta_\epsilon;$ and,

(18) $\Delta u = \dfrac{v'(\epsilon)}{\theta_\epsilon}.$

We can take the partial derivatives of u_L, and u_H with respect to A to obtain

$$(19) \quad \frac{\partial u_L}{\partial A} = \frac{-\theta_{\epsilon A}(1-\theta)v'(\epsilon) - \theta_\epsilon \theta_A v'(\epsilon)}{\theta_\epsilon^2};$$

$$(20) \quad \frac{\partial u_H}{\partial A} = \frac{\theta_{\epsilon A}\theta v'(\epsilon) - \theta_\epsilon \theta_A v'(\epsilon)}{\theta_\epsilon^2}.$$

Equations (19) and (20) indicate how investment alters transfers to the manager in each state of the world. The results are summarized in Proposition 2.

Proposition 2: *If* $\theta_{\epsilon A}(\epsilon, A) \geq 0$, *then* $\partial u_L/\partial A < 0$. *If* $\theta_{\epsilon A}(\epsilon, A) \leq 0$, *then* $\partial u_H/\partial A < 0$. *If* $|\theta_{\epsilon A}| \leq \theta_\epsilon \theta_A$, *then both* $\partial u_L/\partial A < 0$, *and* $\partial u_H/\partial A < 0$.

Proof: The signs of (19) and (20) depend on the sign of $\theta_{\epsilon A}(\epsilon, A)$. If $\theta_{\epsilon A}(\epsilon, A) \geq 0$, then $\partial u_L/\partial A < 0$. The sign of (20) is indeterminate, however, $|\theta_{\epsilon A}| \leq \theta_\epsilon \theta_A$ is sufficient for $\partial u_H/\partial A < 0$. If $\theta_{\epsilon A}(\epsilon, A) \leq 0$, then $\partial u_H/\partial A < 0$. In this case, the sign of (19) is indeterminate, however, $|\theta_{\epsilon A}| \leq \theta_\epsilon \theta_A$ is sufficient for $\partial u_L/\partial A < 0$. *Q.E.D.*

Proposition 2 indicates that if investment and effort are complements (resp. substitutes), then the wage in the low cost (resp. high cost) state falls as investment increases. This is because investment increases (resp. decreases) the marginal product of effort. If the cross-effect of investment on the marginal product of effort is low relative to the marginal products of effort and investment, then wages in both states fall. Investment also alters the incentive constraint facing the manager. The difference Δu is the power of the incentive required to satisfy the incentive constraint and generate the optimal level of managerial effort. In proposition 3 we examine the effect of investment on the power of the incentive.

Proposition 3: *Given A1 and A2, the power of the incentive increases with investment when* $\theta_{\epsilon A}(\epsilon,A) < 0$, *(effort and investment are substitutes), and decreases when* $\theta_{\epsilon A}(\epsilon,A) > 0$, *(effort and investment are complements).*

Proof: From (18) $\Delta u = v'(\epsilon)/\theta_\epsilon$. Therefore, $\partial\Delta u/\partial A = -\theta_{\epsilon A} v'(\epsilon)/\theta_\epsilon^2$. Given *A1* and *A2*, $\theta_\epsilon > 0$, and $v'(\epsilon) > 0$, and the result follows. *Q.E.D.*

The intuition behind this result is straightforward. If $\theta_{\epsilon A} > 0$, then investment increases the probability of the low cost outcome, and increases the marginal productivity of effort for a given incentive scheme. Since $v'(\epsilon)$ is independent of the power of the incentive, the firm can induce the same level of effort with a lower incentive. Moreover, because $\Delta u < 0$, there is a Pareto improvement in risk sharing between the firm and the manager. By analogy, when $\theta_{\epsilon A} < 0$, an increase in investment leads to an increase in the incentive and a worsening of risk sharing. Furthermore, the last term in (15) is positive, which reduces the marginal benefit of an extra unit of investment. This is because when $\theta_{\epsilon A} < 0$, an extra unit of investment increases the power of the incentive, which is costly to the firm because the manager is risk averse.

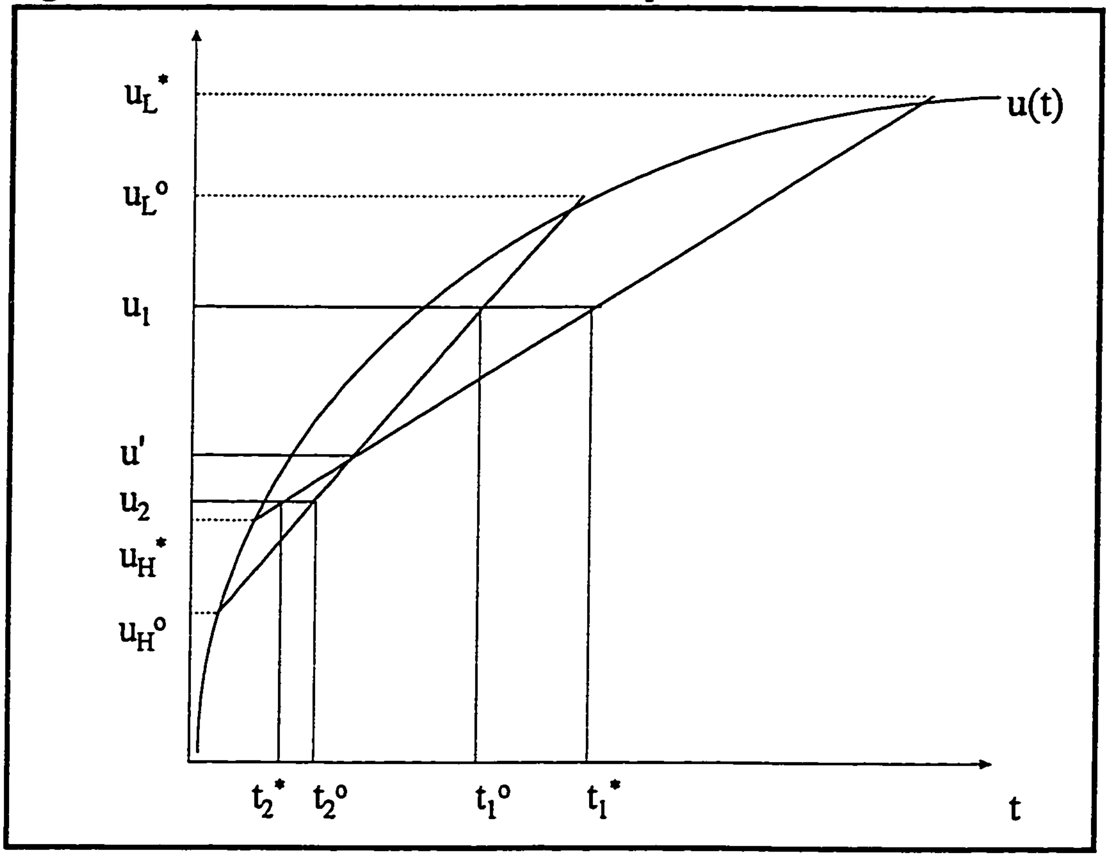
2.5 Example $\theta_{\epsilon A} = 0$, $\epsilon = \epsilon^*$

We now consider an example which illustrates how an increase in investment can lower the cost of implementing a fixed level of effort, ϵ^*. Propositions 2 and 3 show the relationship between investment and the cost of managerial effort. In general, the net effect of investment on the cost of effort is ambiguous. **Figure 1** illustrates the impact of investment on expected transfers when $\theta_{\epsilon A} = 0$. Since the participation constraint is binding, the expected utility from income is fixed at $\bar{u} = \bar{U} + v(\epsilon^*)$. Let u_L^* and u_H^* represent the state dependent utilities consistent with satisfying the participation constraint given $A = A^*$. Consider a discrete increase in A, from A^* to A^o. Since $\theta_{\epsilon A} = 0$, the power of the incentive is unchanged, holding effort fixed, and u_L^* and u_H^* fall by equal amounts to u_L^o and u_H^o. The expected utility function of the manager under (u_H^*, u_L^*) intersects the expected utility function under (u_H^o, u_L^o) at u'. On the *x-axis*, we can measure the change in the expected transfer to the manager due to a change in A.

There are two possible outcomes, depending on whether $\bar{u}$ is greater or less than u'. Let $\bar{u} > u'$. This case is represented graphically by $\bar{u} = u_1$. The expected transfer with $A = A^*$ is given by $t_1^* = \theta(\epsilon^*, A^*) t_{L1}^* + (1 - \theta(\epsilon^*, A^*)) t_{H1}^*$. Because the participation constraint is binding, the manager receives u_1 for all A. Therefore, the increase in investment causes the expected payment to decrease to $t_1^o = \theta(\epsilon^*, A^o) t_{L1}^o + (1 - \theta(\epsilon^*, A^o)) t_{H1}^o$. On the other hand, let $\bar{u} < u'$. This case is represented by $\bar{u} = u_2$. In this case, expected payments increase from $t_2^* = \theta(\epsilon^*, A^*) t_{L2}^* + (1 - \theta(\epsilon^*, A^*)) t_{H2}^*$ to $t_2^o = \theta(\epsilon^*, A^o) t_{L1}^o + (1 - \theta(\epsilon^*, A^o)) t_{H2}^o$.

probability of a low cost outcome. This suggests a relationship between the degree of risk aversion and the probability of a low cost outcome in determining whether or not the firm wishes to increase or decrease A.

Figure 2.1 The Effect of Investment on Expected Transfers when $\epsilon = \epsilon^*$



3. Simulation

In the previous section we showed that, when there is moral hazard, $\tilde{A} \neq A^*$, and, $\tilde{\epsilon} \neq \epsilon^*$ in general. In addition, the sign of $\tilde{A} - A^*$ depends on the interaction of three effects: the substitution of investment for effort as effort becomes more expensive; the decrease in output which results because the cost of producing output increases; and the role of investment in increasing or decreasing the costs of implementing a particular level of effort. In this section we quantify the changes in costs, effort, and investment due to moral hazard through numerical simulation. Due to data limitations and the level of abstraction in the model, the results of the simulations should be viewed as benchmark approximations which could be refined in further research.

3.1 Parameterization of the Model

In choosing parameters for the model we must first choose functional forms, and then assign parameter values. Parameter values have been obtained through the literature wherever possible. Where not possible, sensitivity analysis has been done to determine whether the results are robust across parameter changes. While not all of the results are reported here, we report cases where the model is sensitive to parameter changes.

The probability distribution function $\theta(\epsilon, A)$ is defined as follows. Assume that there exists some intermediate output y which is a function of ϵ and A such that $\theta(y) = \frac{\theta_0 + y}{1 + \theta_1 y}$, where $\theta_0 \in [0,1]$ and $\theta_1 \geq 1$.[12] Let y be determined by the CES production function, $y = y_0[a\epsilon^{\rho} + bA^{\rho}]^{1/\rho}$, where $-\infty < \rho \leq 1$ and $a, b \in (0, 1)$.

The manager's wage utility function is assumed to be characterized by constant

[12] When $\theta_0 > 0$, there is a positive probability of the low cost state being realized with no effort or investment. The assumption $\theta_1 > 1$ implies $\theta(y)$ has an upper bound of $1/\theta_1$ and the low cost state cannot be guaranteed for any level of effort or investment.

relative risk aversion (CRRA), and is written $u(t) = \frac{t^{1-R}}{1-R}$ for $R \neq 1$, and $u(t) = \ln t$ for $R = 1$.[13] We assume that the manager receives a reservation wage of t^R=\$1,000,000 annually.[14] The reservation wage is substituted into the utility function to solve for the reservation wage utility of the manager. The disutility from effort is given by the convex function $v(\epsilon) = g\epsilon^{v}$, where g>0, and $v > 1$.[15] The reservation utility of the manager is interpreted as the wage utility that the manager would receive by expending an average level of effort annually. One can think of this average level of effort as consisting of the contractible elements of effort, such as showing up for work, meeting deadlines, etc.. For simplicity we normalize the average level of effort to zero. Therefore we can think of positive effort in our model as the non-contractible components of effort which contribute to productivity. That is, in our model, though the manager must work to earn the reservation wage, we assume that the reservation utility is defined gross of effort over and above the observable contractible elements of effort. It is assumed that there is increasing disutility from effort. This is consistent with the prevalence of double and triple overtime in many industries.

In order to parameterize abatement costs for the firm, we used the Tennessee Valley

---

[13] In an experimental study, Levy [15] finds support for the hypothesis that investors reveal decreasing absolute risk aversion, and rejects the hypothesis that investors reveal increasing relative risk aversion. These results support the use of the CRRA wage utility function for this analysis.

[14] This reservation wage is consistent with managerial compensation reported in the literature. According to Joskow et al. [10], the average total compensation in 1991 dollars received by CEOs over the period 1972-1990 was \$1,005,450.

[15] The results are qualitatively the same for $v = 1$.

Authority (TVA) as a representative utility. Based on their estimate of expected compliance costs for Phase I, c_L and c_H were set to \$30 and \$50 million respectively.[16] Finally, the cost of capital is assumed to be 11.5% per year, the rate used by Ellerman et al., (1997).

3.2 Results

Table 2.1 shows compliance strategies for the Cobb Douglas production function ($\rho = 0$). We assumed that the distributive shares of effort and investment in the production of y (a and b respectively) are equal. The disutility of effort is captured by the parameters g, which scales the effort variable, and v, which determines the curvature of the disutility function. In the absence of any literature which supports a particular choice of g, we have normalized g=1 for all of the simulations. The scale of effort is in annual units, and is commensurable with the other costs in the model. We computed results for v=2 (quadratic disutility), and v=3 (cubic disutility).[17] In order to capture the impact of increasing the coefficient of relative risk aversion, we solved the model for R=.5, R=1, and R=2.[18]

The results from the Cobb-Douglas simulation support the hypothesis that moral

[16] In the TVA 1995 Annual Report [20], it is estimated that total annual abatement costs for Phase I will fall between \$30 million and \$70 million. Costs in the high cost state are set to \$50 million rather than \$70 million in order to guarantee that the model generates an interior solution to the problem of choosing effort in the second-best case.

[17] The results are qualitatively the same when v=1. Note that the results for v=2 and v=3 can not be directly compared, since the units of effort are not directly commensurable unless we vary g as well.

[18] Newbery and Stiglitz [17] report that the coefficient of relative risk aversion for agricultural producers ranges from R=.3 to R=2.

hazard leads to higher total abatement costs, and the substitution of investment in abatement capital for effort in the firm's compliance strategy. The increase in total annual abatement costs due to moral hazard ranges from 2.49 to 4.08%. In all cases, investment increases while effort and payments to effort decrease. The decrease in effort in percentage terms greatly outweighs the increase in investment which results in a decrease in θ for all cases (-44.87% to -57.38% compared to 9.88% to 18.25% respectively). Finally we can see the effect of the increase in risk aversion on the substitution of investment for effort in the firm's compliance strategy. As R increases, both the percentage decrease in effort and the percentage increase in investment rise relative to the first-best case.

Table 2.1 **Optimal Compliance Strategy and Abatement Costs With and Without Moral Hazard: The Cobb-Douglas Case***

	$R=.5$		$R=1$		$R=2$	
First-best	v=2	v=3	v=2	v=3	v=2	v=3
C^* ($ million)	37.95	37.78	38.07	37.84	38.28	37.95
ϵ^*	0.81	0.78	0.72	0.73	0.61	0.67
A^* ($ million)	15.06	15.18	15.42	15.37	16.00	15.68
θ^*	0.78	0.77	0.77	0.77	0.76	0.76
t^* ($ million)	1.76	1.53	1.69	1.49	1.58	1.42
Second-best						
$\bar{C}$ ($ million)	39.16	38.72	39.42	38.88	39.84	39.12
$\bar{\epsilon}$	0.38	0.43	0.34	0.40	0.26	0.35
$\bar{A}$ ($ million)	16.88	16.68	17.77	17.21	18.92	17.88
$\bar{t}_L$ ($ million)	2.17	1.86	2.11	1.79	1.89	1.64
$\bar{t}_H$ ($ million)	0.003	0.03	0.24	0.27	0.58	0.55
$\bar{\theta}$	0.72	0.73	0.71	0.72	0.69	0.71
$E(\bar{t})$ ($ million)	1.56	1.36	1.57	1.37	1.48	1.33
Percent Changes						
ΔC (%)	3.19	2.49	3.55	2.75	4.08	3.08
Δt (%)	-11.36	-11.11	-7.10	-8.05	-6.33	-6.34
ΔA (%)	12.09	9.88	15.24	11.97	18.25	14.03
$\Delta\epsilon$ (%)	-53.09	-44.87	-52.78	-45.21	-57.38	-47.76
$\Delta\theta$ (%)	-7.69	-5.60	-7.79	-6.49	-9.21	-6.58

* $\theta_0=0$, $\theta_1=1$, $y_0=1$, $a=.5$, $b=.5$, $g=1$

As ρ increases, the inputs ϵ and A become increasingly substitutable. To see the effect of varying ρ, recall that as the cost of effort increases, there are two effects. First, as effort becomes more costly there is a tendency to substitute A for ϵ. For a given level of output, this leads to an increase in investment relative to the first-best case. On the other hand, as effort becomes more costly, the cost of producing a unit of y increases which encourages the firm to reduce y and the level of investment relative to the first-best case. The net effect depends on the value of ρ. In the extreme case of $\rho = -\infty$, the production function for y is Leontief, and inputs are used in fixed proportions. In this case, the cost effect dominates and we expect investment to decrease. On the other hand, if $\rho = 1$, inputs are perfect substitutes, and the firm employs exclusively the cheaper input.

In addition to determining the substitution possibilities between effort and investment, the choice of ρ determines the effect of investment on the marginal product of effort. In particular, after some algebraic manipulation it is possible to show,

$$(18) \qquad \theta_{\epsilon A} \overset{s}{=} \left[(1-\rho)y^{\rho}\frac{dy}{dA}\right]\frac{1}{(1+\theta_1 y)^2} - 2\theta_1 y^{1-\rho}\frac{1}{(1+\theta_1 y)^3}\frac{dy}{dA}.$$

Therefore, a sufficient (but not necessary) condition for $\theta_{\epsilon A} < 0$ is $\rho \geq 1$, while a sufficient (but not necessary) condition for $\theta_{\epsilon A} > 0$ is $\rho \leq -1$. When $-1 < \rho < 1$, the value of $\theta_{\epsilon A}$ is indeterminate and is decreasing in y.

Note that when $\rho = 1$, effort and investment are perfect substitutes in the production of y. Therefore production theory suggests that y be produced with the cheapest input and we would predict complete substitution of investment for effort if effort becomes more expensive than investment. However, in this case $\theta_{\epsilon A} < 0$. Therefore, from Proposition 3 we can see that the incentive constraint introduces a countervailing effect on the increase in investment predicted from production theory. That is, as investment increases, the power of the incentive increases. This decreases the marginal expected profit from increasing A. Similarly, when $\rho \leq -1$, effort and investment are less substitutable in the production of y. As the cost of effort increases, we expect investment to fall through the output effect. Again, Proposition 3 suggests that the incentive constraint introduces a countervailing effect on the decrease in investment predicted by production theory. That is, as investment falls, the power of the incentive decreases, which increases the marginal expected profit from increasing A. Finally, when $-1 < \rho < 1$, $\theta_{\epsilon A} < 0$ for large values of y. Therefore, for large values of y, the power of the incentive will be increasing in A, and the marginal expected profit from increasing A will fall due to the increased costs of the incentive constraint.

We assume that investment and effort are neither perfect substitutes, nor perfect complements. In Table 2 we show the effect of varying the input elasticity of substitution (σ) for intermediate cases when R=1. We solved the model for $\rho = -1$ ($\sigma = .5$), the Cobb-Douglas case $\rho = 0$ ($\sigma = 1$), and $\rho = .5$ ($\sigma = 2$). Again, total abatement costs are always higher in the second-best case. The expected wage bill and θ both decrease relative to the first-best case. As expected, percentage changes in effort and investment are sensitive to the elasticity of substitution. We see that when $\rho = 0$, and $\rho = .5$, investment increases and effort decreases.

However, when $\rho = -1$, we see that *both* investment and effort decrease. The percentage increase in costs falls as the elasticity of substitution rises. This is to be expected, since as ρ increases there is more flexibility for the firm to substitute towards the cheaper input. Finally, investment increases and decreases according to our predictions from production theory with respect to the elasticity of substitution.

Table 2.2 Optimal Compliance Strategy and Abatement Costs With Varying Input Elasticity of Substitution*

	$\rho=-1$		$\rho=0$		$\rho=.5$	
First-best	v=2	v=3	v=2	v=3	v=2	v=3
C^* ($ million)	40.72	40.55	38.07	37.84	36.72	36.01
ϵ^*	0.93	0.88	0.72	0.73	0.36	0.47
A^* ($ million)	5.74	5.64	15.42	15.37	18.60	18.13
θ^*	0.62	0.60	0.77	0.77	0.86	0.86
t^* ($ million)	2.36	2.00	1.69	1.49	1.14	1.11
Second-best						
$\bar{C}$ ($ million)	43.22	42.57	39.42	38.88	36.41	36.25
$\bar{\epsilon}$	0.49	0.53	0.34	0.40	0.05	0.14
$\bar{A}$ ($ million)	5.32	5.19	17.77	17.21	20.59	19.67
$\bar{t}_L$ ($ million)	3.93	3.11	2.11	1.79	1.14	1.14
$\bar{t}_H$ ($ million)	0.46	0.45	0.24	0.27	0.47	0.47
$\bar{\theta}$	0.47	0.49	0.71	0.72	0.85	0.85
$E(\bar{t})$ ($ million)	2.11	1.75	1.57	1.37	1.04	1.04
Percent Changes						
ΔC (%)	6.14	4.98	3.55	2.75	0.80	0.67
Δt (%)	-10.59	-12.50	-7.10	-8.05	-8.77	-6.30
ΔA (%)	-7.32	-7.98	15.24	11.97	10.70	8.49
$\Delta\epsilon$	-47.31	-39.77	-52.78	-45.21	-86.11	-70.21
$\Delta\theta$ (%)	-24.19	-18.33	-7.79	-6.49	-1.16	-1.16

* $\theta_0=0$, $\theta_1=1$, $y_0=1$, $a=.5$, $b=.5$, $g=1$, R=1

4. Conclusions

The moral hazard model and simulation results presented in this paper suggest the type of distortions in compliance strategies we might observe from electric utilities under the CAAA. Bohi and Burtraw [2] argue that, while the market has achieved significant gains in reducing compliance costs, that cost savings are still less than their potential due to idiosyncrasies in the program design which have promoted uneconomic investments and prevented firms from realizing additional savings. Proposition 1 shows that we can expect firms which have moral hazard problems to have higher expected abatement costs than firms in which they don't exist. Dalen and Gomez-Lobo [6] also show that traditional cost function estimates produce biased results if moral hazard is present. In particular, scale economies are overestimated.

In our simulation, cost increases due to moral hazard ranged from 0.91% to 11.08%. On the other hand, an increase in investment will lead to a reduction in permit prices, since holding permits and investing in abatement capital are substitute abatement strategies [7]. In our model, investment rises when effort and investment are highly substitutable. In this case, permit prices will be lower than expected. This result is consistent with, and may explain part of the discrepancy between, actual and predicted compliance costs and strategies observed in the first year of Phase I (see [2, 7]).

Propositions 2 and 3 show that firms use investment to reduce the costs of the incentive and participation constraints on managerial effort. The model predicts that seemingly identical firms will choose different compliance strategies solely because of differences in levels of managerial risk aversion. This may explain the variation in the volume of trading activity across firms observed by Bohi and Burtraw [2].

One might ask whether there is some regulatory policy that could eliminate the distortion in the investment decision caused by moral hazard. We argue that this is not the case, since if managerial effort is not observable, there is no reason to expect that a regulatory agency could do better than the firm in minimizing costs. In our model, the shadow cost of public funds is implicitly zero. If public funds are costly, and revenue from permits is used to balance the budget, overinvestment in scrubbers imposes an additional cost by driving down the price of permits and decreasing public revenues.

The model presented here is greatly simplified and abstracts from many factors that influence utility behaviour. Nonetheless, the results appear to be consistent with stylized facts about the performance of the CAAA permit market in the initial stages of Phase I. This demonstrates the need to treat the firm as more than a black box when predicting compliance costs and strategies. Further research is necessary to determine the actual technological and behavioural relationships between effort and investment in reducing abatement costs.

5. References

[1] D. R. Bohi and D. Burtraw, Utility investment behaviour and the emission trading market, *Resour. Energy* **14**, 129-153 (1992).

[2] D.R. Bohi and D. Burtraw, SO_2 allowance trading: how experience and expectations measure up, Resources for the Future, Discussion Paper 97-24 (1997).

[3] G. Bose, Interlinked contracts and moral hazard in investment, *Journal of Development Economics* **41**, 247-73 (1993).

[4] S. Bryan and L. Hwang, CEO compensation in a regulatory environment: an analysis of the electric utility industry, *Journal of Accounting, Auditing, and Finance* **12**, 223-51 (1997).

[5] J. S. Coggins and V. H. Smith, Some welfare effects of emission allowance trading in a twice regulated industry, *J. Environ. Econom. Management* **25**, 275-297 (1993).

[6] D.M. Dalen and A.Gomez-Lobo, Estimating cost functions in regulated industries characterized by asymmetric information, *European Economic Review* **41**, 935-42 (1997).

[7] A. D. Ellerman, R. Schmalensee, P.L. Joskow, J. P. Montero, and E.M. Bailey, Emissions trading under the U.S. Acid Rain Program: evaluation of compliance costs and allowance market performance, MIT, Center for Energy and Environmental Policy Research, Book # 23 (1997).

[8] S. J. Grossman and O. D. Hart, An analysis of the principal-agent problem, *Econometrica* **51**, 7-45 (1983).

[9] R. W. Hahn and C. A. May, The behaviour of the allowance market: theory and evidence, *Electricity J.* March, 29-37 (1994).

[10] P. Joskow, N. Rose, and A. Shepard, Regulatory constraints on CEO compensation, Brookings Papers: Microeconomics (1993).

[11] P. L. Joskow, N. Rose, and C. Wolfram, Political constraints on executive compensation: evidence from the electric utility industry, *Rand Journal of Economics* **27**, 165-82 (1996).

[12] P. W. Kennedy, Learning about environmental damage: implications for emissions trading, Discussion Paper 97-2, Department of Economics, University of Victoria, Victoria, BC (1997).

[13] J-J Laffont and J. Tirole, "A Theory of Incentives in Procurement and Regulation", The MIT Press, Cambridge MA (1993).

[14] J-J. Laffont and J. Tirole, Pollution permits and compliance strategies, *Journal of Public Economics* **62**, 85-125 (1996).

[15] H. Levy, Absolute and relative risk aversion: an experimental study, *Journal of Risk and Uncertainty* **8**, 289-308 (1994).

[16] W. D. Montgomery, Markets in licenses and efficient pollution control programs, *J. Econom. Theory* **5**, 395-418 (1972).

[17] D.M.G. Newbery, and J.E. Stiglitz, "The Theory of Commodity Price Stabilization", Clarendon Press, Oxford, UK (1981).

[18] E.T. Nwaeze and J.R. Mereba, Market implications of regulatory form in the electric utility industry: an assessment of incentive regulation, *Journal of Accounting, Auditing, and Finance* **12**, 285-307 (1997).

[19] C. Shapiro, Investment, moral hazard, and occupational licensing, *Review of Economic Studies* **53**, 843-62 (1986).

[20] Tennessee Valley Authority, *Annual Report* (1995).

[21] J. Thomas, Discussion: CEO compensation in a regulatory environment: an analysis of the electric utility industry, *Journal of Accounting, Auditing, and Finance* **12**, 252-55 (1997).

[22] J. T. Tschirhart, Transferable discharge permits and profit-maximizing behaviour, in "Economic Perspectives on Acid Deposition Control" (T.D. Crocker, Ed.), Butterworth, Boston, MA (1984).

[23] M. L. Wald, Risk-shy utilities avoid trading emission credits, *Wall Street Journal*, Jan. 24. (1992).

Essay 3

Markets for Water Rights Under Environmental Constraints

1. Introduction

The debate over the merits of alternative allocation mechanisms for surface water rights is increasingly important given persistent water shortages in the Western United States and Canada and increasing demands on existing fresh water sources for hydro-electricity, new industry, and irrigation [23]. Many water sources are fully allocated and there are pressures to reallocate water to suit changing economic conditions. There are limited avenues available to meet new demand. These include supply augmentation, a decrease in existing entitlements, or voluntary transfers [25]. Voluntary transfers are attractive for many reasons; politicians are not seen to arbitrarily assign winners and losers in the allocation process, and water is diverted to its highest benefit use [4].

The allocation problem is compounded by environmental concerns. Minimum instream flows are required to maintain wildlife and recreation values. At the same time, pollution from irrigation as well as industrial and municipal sources is decreasing the quality of the resource. In this paper we examine the efficiency of a market for tradable water rights in a dendritic water system when there are environmental constraints. The results from our paper show that the regulator has limited options for controlling environmental externalities when water is allocated through a market.

Efficiency gains from tradable water rights appear to be significant. Wong and Eheart

[27] find that market systems recoup about 95 percent of the economic value of the optimal distribution of water. This is because the market equates marginal benefits between users. Wong and Eheart's results are derived for a lentic (lakelike) system. The same results do not hold for dendritic systems where the total amount of water available for diversion depends on the upstream return flow regime.[1] In this case it is generally not optimal to equate the marginal benefits of water consumption. Upstream water diversion introduces a positive externality because return flows can be reused downstream. For example, Burness and Quirk [5] show that the total amount of water available for use is maximized by allocating all of the water available for diversion to the first user, and all of the return flows to subsequent users. On the other hand, reductions in return flows due to upstream consumption (as opposed to diversion) result in third party damages when downstream users depend on return flows to exercise their consumptive rights.

The existence of third party damages is related in part to the definition of the water right associated with the license. Water rights are typically defined either in terms of consumptive use or diversion. Hartman and Seastone [9] show that when water rights are defined in terms of consumptive use rather than diversion, markets for water rights are efficient. This is because third party effects from transfers are eliminated when the total amount of water allocated for consumption is less than or equal to the total flow in the system. However, since return flows can be reused downstream, the total amount of water

---

[1] Return flows are defined as the difference between the amount of water diverted for use and the amount of water actually consumed. Return flows are usually quantified by the fraction of water returned to the system.

available for allocation is greater than the total flow in the system. Therefore some observers have argued that a second market for return flows should also be established (see [9, 20]).

Johnson et al. [12] show that third party effects can arise even when the water is defined in terms of consumptive use. As long as return flows are positive, it is necessary to divert more water than allocated under the consumptive right. Physically binding flow constraints between users can arise when the level of water diversion required to satisfy the right exceeds the streamflow. In these circumstances, it is not optimal, in general, for marginal benefits to be equal across users.

One problem which has not been dealt with in allocation models for surface water rights is the relationship between water consumption and water quality. Water quality is typically measured as a concentration of pollutants, and therefore depends on water flows. Changes in the spatial distribution of water use can decrease the capacity of a river system to handle pollutants downstream. Although many authors have acknowledged the relationship between water flows and water quality (see [7, 24]), there is no model which incorporates this relationship into an optimal allocation rule for surface water rights. Reductions in water quality degrade both private benefits from consumptive water use, as well as private and public benefits from non-use values. Because of the interaction between water flows, pollutant loadings, and water quality, it is necessary to specify the allocation of pollution rights in any model of surface water allocation.

Tradable permit systems for allocating pollutants in watersheds are slowly being implemented [1].[2] Montgomery [19] shows that the efficiency of a market system for allocating pollutants depends on whether permits are defined in terms of emissions or damages. In general, permit systems defined in terms of emissions are not efficient since they don't capture the spatial characteristics of damages from firms [19, 18]. Modelers get around the spatial problem by assuming that marginal social damages from pollution are identical for all firms. However this assumption is troublesome when applied to the problem of water quality. In particular, impacts from pollutants are not equal for upstream and downstream users. As pollutants accumulate and water flows decrease, marginal damages from downstream firms increase. The cumulative effect creates a technological externality from upstream polluters because input choices for downstream users are restricted.

Unger [26] investigates whether decentralized decision making can lead to an efficient allocation of pollutants along a river. In Unger's model, firms produce heterogeneous effluents and the regulator must consider the way chemicals entering upstream interact with chemicals entering downstream in producing aggregate water quality. If the aggregate water quality production function is completely recursive [3], then socially

---

[2] The EPA's policy is to support and promote the use of tradable permits for effluent discharges in order to maintain water quality standards under the Clean Water Act. This has promoted the development of experimental trading programs in several watersheds.

[3] Complete recursiveness is a technological assumption on the production function which permits sequential separability in decision making. This framework is applicable to problems where there are directional externalities, as in the water pollution problem considered by Unger [26]. See Blackorby et al. [2] for a discussion of asymmetric separability, homothetically recursive functions, and recursively decentralized decision making.

optimal input mixes can be achieved by assigning responsibility for cumulative water quality to each firm on the river. This property of the technology allows the water agency to derive location specific imputed price aggregates for input use at each site. Using these imputed social values, the agency can force firms to bear the full social cost of input use, and it is not necessary for the authority to assign input vectors directly to each firm. Unger's model assumes that emissions and their effects are linked by fixed delivery ratios, that is, damages are proportional to discharges. However, Braden et al. [3] show that changing the delivery ratio of pollutants is a substitute for reducing emissions which needs to be considered for maximum efficiency in meeting water quality objectives.[4]

Given these results, a market for water rights presents two management problems. First, trading of consumptive rights can occur between non-adjacent users which generates third party effects as the return flow regime is altered for users in between them. Second, changes in the surface water regime will result in changes to the delivery ratio of pollutants. One implication of Unger's model is that a market must be capable of supporting multiple location-specific prices. However, permits entitle users to a homogeneous good (the right to consume water). The question of existence of equilibrium in a market with several prices is complicated by the possibility of arbitrage between users. Finally, reductions in water use and reductions in emissions are substitute strategies for maintaining water quality. Therefore we must design a regulation which generates the optimal substitution of environmental

---

[4] In Braden et al. [3], altering the delivery ratio means altering the transportation regime for pollutants. In this context the authors consider the problem of optimal containment of pollutants as an alternative strategy to reducing emissions.

inputs along the stream.

In this paper we develop a general model which is used to characterize the optimal distribution of consumptive water use and pollution loadings along a river. In order to capture competing management objectives we assume that there are minimum instream flow and quality requirements at any given location on the river. We also assume that there are compact agreements which specify the flow and quality of water which must be delivered at the jurisdictional boundary of the watershed. The model applies to point sources of pollution where damages can be attributed to particular users.[5]

We examine the efficiency of alternative government policies for maintaining water quality when consumptive water rights are allocated by a market mechanism. We find that when there is a single price equilibrium for consumptive water rights, there is no policy which efficiently allocates consumptive water rights and pollutants given water quality objectives. We argue that a single price equilibrium is a result of myopia on the part of water users. We show that the marginal benefit of purchasing surface water rights depends on the location of the seller. In particular, if licenses are purchased from upstream users, both water quality and water flows are increased for the buyer. Purchasing licenses from downstream users does not generate any similar externality. This effectively divides the market into an upstream and a downstream segment at each site. In this case, a market for

[5] Point source models are often used to describe situations where there are large industrial polluters along a river. Our model also applies to agricultural users located along the river. Xepapadeas [28] shows that if emissions are partially observable, it is possible to turn non-point source problems into point source problems.

surface water rights coupled with a Montgomery style market to allocate a fixed level of pollution in the watershed generates the optimal outcome. A Nash equilibrium in both markets supports location specific pricing of permits.

The remainder of the paper proceeds as follows. In section 2 we solve for the socially optimal allocation of surface water and effluent discharges along a stream. In section 3 we examine the policy implications. In the main proposition of the paper we show that a Nash equilibrium in markets for both consumptive water and tradable pollution rights can support socially optimal location-specific prices. We conclude by discussing the feasibility of the solution, and suggest ideas for future research.

2. Optimal Allocation of Water and Pollutants in a Dendritic System

2.1 Characterization of Streamflows and Benefits

2.1.a Streamflows

Let N represent an ordered set of n water users located along a river system. There are no branches feeding into the river system. The flow of water between each user thus depends only on the water flow at the source, and the allocations to and return flows from, prior users. Users are ordered $i=1,\ldots,n$, in order of increasing distance from the source. The total amount of water available for consumption at the source is denoted by v_o. The amount of water $v(i)$ available for diversion for the ith user depends on the consumption of water by upstream users. The total amount of water consumed by user i is equal to the amount diverted less the return flow,

$c(i) = (1 - R^i)s(i),$

where $s(i)$ is the amount of water allocated for consumption at location i, and $c(i)$ is the amount of water consumed.[6] The return flow function for each user is exogenous and given by the parameter R^i.[7] Therefore the amount of water available for diversion at any point along the river evolves according to the first order difference equation,

$v(i+1) - v(i) = -(1 - R^i)s(i)$.[8]

Similarly, water quality at any reach downstream from a specified point source of pollution is determined by a first order difference equation [24]. Let $q(i)$ denote the level of water quality at user i's intake and $e(i)$ denote the amount of effluent that user i chooses to discharge. For simplicity assume that the level of effluent produced by user i can be captured by a single dimension variable e which is a monotonically increasing function of the level of effluent-generating inputs.The difference in water quality for users i and $i+1$ can then be written

$q(i+1) - q(i) = f^i(c(i), e(i), v(i), q(i)).$

---

[6] Although the stochastic nature of flows along the river has important implications for the quality and quantity of water available for consumption, we assume that water flows are deterministic in order to highlight the problems which arise from separating the regulation of water consumption from water quality. If a particular regime cannot support an optimum in a deterministic setting, then the regime also cannot support it in a stochastic setting.

[7] Burness and Quirk [5] provide ranges for return flow parameters associated with particular water uses. These **R** values range from 5-10% for evaporative cooling, to 30-60% for agricultural use, and 80-90% for domestic and municipal use.

[8] Note that the superscript i denotes variables or functional forms which are parametric to user i and also appears as an argument of user i's choice variables because the optimal solution will depend on user i's location along the river.

The difference in water quality depends on the amount of water consumed and effluent discharged by i. It is assumed that quality falls with water consumption and effluent discharge so that $f_c^i < 0$ and $f_e^i < 0$. Streamflows and the level of water quality are assumed to affect the change in quality in two ways. First, the negative effects of decreased flows and increased levels of pollution are cumulative. That is, as the concentration of pollutants increases, greater reductions in water quality result from the addition of extra units of effluent, or reductions in available flows.[9] Second, the assimilative capacity of the stream is assumed to be lower at lower levels of quality and water flows so that $f_q^i > 0$ and $f_v^i > 0$. Positive assimilative capacity of the stream implies that water quality improves downstream ceteris paribus. Therefore $f^i(0, 0, v(i), q(i)) > 0$ for all $q(i) < q_o$ where q_o indicates quality at the source.[10] Finally, the function $f^i(.)$ may depend on site specific characteristics at i. The level of overall quality at any intake for user i is then

$$q(i) = \sum_{j=1}^{i-1} f^j(c(j), e(j), v(j), q(j)) + q_o.$$ [11]

---

[9] For example, pulp mill effluent treatments create large benefits if dissolved oxygen levels are low; algal growth is stimulated at an exponential rate by temperature increases; and there is a positive correlation between concentrations of total nitrogen or total phosphorous and suspended solids [22].

[10] We are implicitly assuming that water is pristine at the source, so that the initial level of quality can't be improved upon. This ignores the influences of natural sedimentation and other disturbances on water quality. However, since this model is not stochastic, the assumption does not alter our qualitative results.

[11] The level of generality associated with this specification of the quality function allows different methods for defining quality to be incorporated in the model.

2.1.b Characterization of Benefits

Each user has a benefit function which depends positively on the amount of water consumed, as well as on other inputs which may produce effluent. The benefit function for user i may also depend on the quality and quantity of water available at i's intake. For example, low water flows may increase diversion costs, or water of lower quality may increase operating costs and generate lower returns than water of high quality. Finally the benefits of water consumption may depend on location along the river. For example, in the agricultural sector the quality of land may change along the water system, affecting returns per unit of water consumed. Therefore, the benefit function for user i can be written

$B^i = B^i(c(i), e(i), v(i), q(i))$ with

$B_c^i \geq 0$, $B_e^i \geq 0$, $B_v^i \geq 0$, and $B_q^i \geq 0$.

The model allows for heterogeneous users within a sector, as well as heterogeneous industries, to use water along the stream. It is assumed that pollution-producing inputs and water are substitutes.[12] Finally, it is assumed that users must consume water in order to derive positive benefits from a water right. However, there may be non-consumptive users such as recreationists located along the stream who enjoy positive marginal benefits from increases in stream quality and water flows. In order to incorporate non-consumptive uses, it is assumed that there are social benefits derived from increased water quality and flow levels. It is assumed that these benefits are recognized by a regulator who optimally sets

[12] For some industries such as agriculture, these inputs are often modelled as complements. Leontief benefit functions result in allocations which are corner solutions. Since the purpose of the model is to capture the marginal impacts of water reallocation on users in the system, we have ignored the case of Leontief benefit functions.

instream flow constraints as discussed below.

2.1.c Environmental and Contractual Constraints.

Two types of constraints are assumed to be imposed: endpoint and instream flow needs. A compact agreement stipulates that a minimum quantity of water $\bar{v}$ and level of quality $\bar{q}$ be available downstream from the last (n^{th}) user. Instream flow need (IFN) constraints are determined by the regulator and specify minimum levels of quality, $\tilde{q}$, and flows, $\tilde{v}$, which must be maintained within the river. It is assumed that $\tilde{v} < \bar{v}$, since otherwise $\bar{v}$ would not bind. To satisfy user *i*'s consumptive rights, the amount of water available at *i*'s intake must satisfy the constraint

$$v(i) - \tilde{v} \geq c(i)/(1 - R^{i}).$$

If $\tilde{v} = 0$, this constraint is equivalent to the flow constraint between users that appears in Johnson et al. [12]. Third party effects from water transfers result when this constraint is binding for some user.

The IFN constraint on quality, $\tilde{q}$, may be greater or less than $\bar{q}$. If $\tilde{q} > \bar{q}$, then it is the IFN constraints on quality that will be binding, and the end point quality constraint will always be satisfied. If $\tilde{q} < \bar{q}$, then the endpoint quality constraint will dictate cleanup unless the assimilative capacity of the stream is sufficiently large.

2.2 The Regulatory Problem

The regulatory problem is to maximize the value of water consumption subject to the streamflow constraints specified above. This problem is modelled as an optimal control problem where the regulator chooses the optimal allocation of water, $s(i)$, and pollution, $e(i)$, for each user i along the river. The set of admissible controls, W, is defined by the IFN constraints on water flows and quality as well as by non-negativity constraints on the control variables,

$$W(v(i),q(i),i) \equiv \{s(i),e(i) \mid g^1 = v(i) - s(i) - \bar{v} \geq 0;$$
$$g^2 = q(i+1) - \tilde{q} \geq 0;\ c(i) = (1-R^i)s(i);\ s(i) \geq 0;\ e(i) \geq 0\}$$

The first constraint, g^1, is the flow constraint between users. It stipulates that the amount of water allocated can not exceed the amount of water available in the system at that point. As Johnson et al. [12] note, this constraint is more likely to bind when $\bar{v}$ is small relative to v_o. It is less likely to bind when consumptive use is small and there are high return flow coefficients. The second constraint, g^2, requires that water quality at the diversion point for user $i+1$ be at least $\tilde{q}$.

The problem for the regulator is to choose the optimal path along the river of water allocated for diversion, $s^*(i)$, and effluent discharge $e^*(i)$ by solving

$$\text{(P1)} \quad \underset{s(i),e(i)}{Max} \sum_{i=1}^{n} B^i(c(i), e(i), v(i), q(i))$$

subject to

$$\text{(P1.1)} \quad v(i+1) - v(i) = -(1-R^i)s(i) \qquad (\mu_1(i) \geq 0);$$

(P1.2) $q(i+1) - q(i) = f^i(c(i), e(i), v(i), q(i))$ $(\mu_2(i) \geq 0)$;

(P1.3) $g^1 = v(i) - s(i) - \tilde{v} \geq 0$ $(\lambda_1(i) \geq 0)$;

(P1.4) $g^2 = q(i) + f^i(c(i), e(i), v(i), q(i)) - \tilde{q} \geq 0$ $(\lambda_2(i) \geq 0)$;

(P1.5) $v(1) = v_o$; $v(n+1) = \bar{v}$; $q(1) = q_o$; $q(n+1) = \bar{q}$, and

(P1.6) $s(i) \geq 0$; $e(i) \geq 0$.[13]

This is a general constrained control problem with two control variables, s and e, and two state variables v and q. The Lagrangian is

$$\mathcal{L}(c,e,\mu_1,\mu_2,\lambda_1,\lambda_2,i) =$$
$$H(c(i),e(i),v(i),q(i),\mu_1(i),\mu_2(i)) + \lambda_1(i)[g^1(i)] + \lambda_2(i)[g^2(i)],$$

where

$$H(c(i),e(i),v(i),q(i),\mu_1(i),\mu_2(i)) \equiv B^i(c(i),e(i),v(i),q(i))$$
$$+ \mu_1(i)[-(1-R^i)s(i)] + \mu_2(i)[f^i(c(i),e(i),v(i),q(i)]\ ,$$

is the Hamiltonian associated with the problem.

2.3 Characterization of the Optimal Solution

Using the Maximum Principle we can characterize the optimal solution to this problem. The necessary conditions for a maximum are

(1) $$\frac{\partial \mathcal{L}}{\partial s(i)} = B_c^i - \mu_1(i) + \mu_2(i) f^i_{c(i)} - \lambda_1(i)/(1-R^i) + \lambda_2(i) f^i_{c(i)} \leq 0,$$

[13] Note that as long as the benefit functions are monotonically increasing in *e(i)* and *c(i)*, the compact constraints $\bar{q}$ and $\bar{v}$ will be binding.

$$\frac{\partial \mathcal{L}}{\partial s(i)} s(i) = 0 \; ; \quad i = 1,...n.$$

(2) $$\frac{\partial \mathcal{L}}{\partial e(i)} = B_e^{\,i} + \mu_2(i) f_{e(i)}^{i} + \lambda_2(i) f_{e(i)}^{i} \le 0,$$

$$\frac{\partial \mathcal{L}}{\partial e(i)} e(i) = 0 \; ; \quad i = 1,...n.$$

(3) $$\mu_1(i) - \mu_1(i-1) = \Delta\mu_1(i) = -\frac{\partial \mathcal{L}}{\partial v(i)} =$$

$$-[B_v^{\,i} + \lambda_1(i) + (\mu_2(i) + \lambda_2(i)) f_{v(i)}^{i}]; \qquad i = 2,...n.$$

(4) $$\mu_2(i) - \mu_2(i-1) = \Delta\mu_2(i) = -\frac{\partial \mathcal{L}}{\partial q(i)} = -[B_q^{\,i} + \lambda_2(i) + (\mu_2(i) + \lambda_2(i)) f_{q(i)}^{i}]$$

$i = 2,...n.$

(5) $v(i) - s(i) - \bar{v} \ge 0 \; ; \quad \lambda_1(i)[v(i) - s(i) - \bar{v}] = 0 \; ; \quad i = 1,...n.$

(6) $q(i+1) - \bar{q} \ge 0 \; ; \quad \lambda_2(i)[q(i+1) - \bar{q}] = 0 \; ; \quad i = 1,...n.$

(7) $v(n) - (1 - R^{\,i}) s(n) - \bar{v} \ge 0;$

(8) $q(n+1) - \bar{q} \ge 0.$

A sufficient condition for the solution to be a maximum is that the Lagrangian be concave with respect to *c(i)*, *e(i)*, *v(i)*, and *q(i)*. The following assumptions guarantee an interior solution to the control problem:[14]

Assumption 1 (**A1**)*: The benefit function is increasing and strictly concave in each of its arguments, c,e,v and q, for all users i;*

Assumption 2 (**A2**)*: The quality function is decreasing and strictly concave in the arguments c and e, and increasing and concave in its arguments v and q for all users i.*

Assumption 3 (A3): Input consumption is strictly positive for all users, $c(i) > 0, \quad e(i) > 0, \quad \forall\, i.$

Making the appropriate substitutions, we can describe the optimal shadow values of input use at each site:

(9) $\mu_1(i) = B_c^i - (B_e^i/f_e^i)f_c^i - \lambda_1(i)/(1 - R^i);$

(10) $\mu_2(i) = -(B_e^i/f_e^i) - \lambda_2(i);$

(11) $\Delta\mu_1(i) = -[B_v^i + \lambda_1(i) - (B_e^i/f_e^i)f_v^i];$

(12) $\Delta\mu_2(i) = -[B_q^i + \lambda_2(i) - (B_e^i/f_e^i)f_q^i].$

The left hand sides of equations (9) and (10) represent the marginal cost of water consumption and pollution discharge respectively at site *i*, and can be thought of as the shadow value of input use at that site. These shadow values represent the prices attached to

[14] See Leonard and Long [15] for sufficient conditions for a global maximum in a constrained control problem.

input use which support the optimum. From (9) we see that, at the optimum, the marginal social cost of water consumption is set just equal to the *net* marginal benefit of allocating an extra unit of water to user i. This consists of the marginal benefit of consumption (the first term on the RHS), minus the marginal private cost which results because consumption of water increases the marginal damage of an extra unit of effluent, thus constraining user i's choice of pollution producing inputs (the second term on the RHS). Finally, if the IFN constraint becomes binding, this decreases the net benefit of allocating water for consumption at i. Similarly, equation (10) shows that at the optimum the marginal social cost of effluent discharge at site i is just equal to the *net* marginal benefit of the discharge, which is equal to the marginal benefit from using pollution producing inputs weighted by the marginal damage caused from emitting an extra unit of effluent, $-(B_e^i/f_e^i)$.

Equations (11) and (12) show the optimal evolution of shadow prices for using water and effluent producing inputs. Given (A1) and (A2), $\Delta\mu_1(i) < 0$ and $\Delta\mu_2(i) < 0$, showing that the shadow cost of input use declines as one proceeds downstream. This is because upstream water consumption and pollution reduce the value of the state variables, $v(i)$ and $q(i)$, which creates a negative externality for all downstream users. The shadow cost of input use declines downstream as the change in the state variables affects fewer users. Note that if $B_v^i = B_q^i = f_v^i = f_q^i = 0$, and the IFN constraints are not binding, ie., $\lambda_1(i) = \lambda_2(i) = 0$ for all i, then the externalities disappear. The control problem then reduces to one of allocating water and pollutants in a lentic system, and the marginal benefits of input use are equalized across users. This is consistent with the findings of Hartman and Seastone [9], and

Johnson et al. [12]. Thus the problem of allocating water and pollution rights on a lentic system is just a special case of the more general problem of allocation when there are directional spatial externalities as in the model here.

If either of the IFN constraints on water flows or water quality are binding ($\lambda_1(i) \neq 0$ and/or $\lambda_2(i) \neq 0$ for at least one user $i \in N$), then from (11) and (12) we see that for each i at which the IFN constraints bind, there is a decrease in the path of shadow costs attached to downstream input use. This reflects the relative increased cost of input use upstream from i, since shifting input use from upstream to downstream users would increase both the volume and quality of flows at site i, thus reducing i's flow constraints.

3. Market Decentralization

We now address the question of whether a decentralized market for water licenses is efficient when there are water quality constraints. We examine different regulatory options for controlling pollutants and evaluate the performance of each market equilibrium by comparing the resulting allocation of inputs to the allocation which follows from (1)-(8).

The results depend on whether or not users are "myopic". Users can be myopic in the sense that they do not account for the increases in water volume and quality which arise from a redistribution of water rights from upstream users to themselves or further downstream in their decisions to purchase licenses. The benefit of the license is then independent of the location from which it is purchased, resulting in a single price equilibrium.[15] If users are not myopic, they recognize that the value of a license depends on the site from which is purchased. In this case the externalities and third party effects introduced by reallocating rights away from upstream users are incorporated into each user's decision function, and a multiple price market results.

The market for tradable water rights is described as follows. The regulator allocates water licenses $\bar{W} = \sum_{i=1}^{n} w_i^o$ where w_i^o is the initial allocation to each user i, and $\bar{W} = v_o - \bar{v}$ is the total amount of water available for consumption. A water right entitles the user to one unit of water consumption. If $w(i)$ is the number of water licenses held by user i, user i is in compliance when $c(i) \leq w(i)$.[16] This constraint holds with equality in equilibrium as long as license prices are positive.[17] We assume that all of the markets considered below are perfectly competitive in the sense that there are many bidders in the market for each

[15] Given that users are rational, we assume that myopia occurs as a result of institutional design which prevents users from identifying the external benefits attached to the purchase of licenses. For example, buyers and sellers may trade through a central broker rather than directly with each other.

[16] For a discussion of the impacts of non-compliance on permit market outcomes see Malik [17], and Keeler [13].

[17] The result is proved formally in Theorem 3.2, Montgomery [19]. However Montgomery's proof does not automatically hold for consumptive water licenses. One can construct examples in which IFN constraints for water licenses are binding and there are unused licenses in equilibrium. Since the right to consume water is not equivalent to the right to divert water, there may be cases (for example when return flow parameters across users are "high enough") where no user can divert enough water to exhaust consumptive rights without violating the IFN constraint $\bar{v}$. Note that $R^i = 0$ implies that the required diversion is just equal to the consumptive right. If $R^i = 0$, and the marginal benefits of water use are positive for some user i, then that user will pay positive prices for licenses from any location. Therefore no user holds excess licenses in equilibrium.

license.[18]

In Section 3.1 we will examine the efficiency of four different policies for regulating pollution when water rights are tradable and users are myopic. In section 3.2 we consider the case when users are not myopic. We show that in this case it is possible to implement the optimal solution when there is a corresponding market for tradable rights for pollution damages. In Section 4 we discuss the feasibility of designing an institution that could generate an efficient allocation of rights.

3.1 Water License Market Equilibrium with Myopic Users

In this section we examine four alternative approaches to regulating water quality when water licenses are tradable: applying firm specific allowable discharges; maintaining IFN quality constraints at each site; creating markets for discharge licenses; and creating markets for pollution damage licenses. These policies illustrate the range of options available for regulating water quality. Each policy consists of a choice of target variable (emissions versus damages) and a choice of instrument (command and control versus market). Since users are myopic they assume that the only benefit from a water license is the right to consume water. Therefore the value of each water license is independent of the location from which it is purchased and users face a single price in the market.

[18] For a discussion of imperfect competition in permit markets see Hahn [10].

3.1.a Site-specific allowable effluent discharges

Assume that the regulator for water quality approves user effluent discharges at some level $\hat{e}(i)$ according to $q(i+1) \geq \tilde{q}$ based on a given level of water volume flowing to and being consumed by user i. Then each user solves

$$\text{(P2.a)} \quad \underset{c(i),e(i),w(i)}{Max} \quad B^i(c(i),e(i),v(i),q(i)) - p_w(w(i) - w_i^o)$$

subject to

$$\text{(P2.1)} \qquad w(i) - c(i) \geq 0; \qquad\qquad (\alpha_1(i) \geq 0);$$

$$\text{(P2.2)} \qquad \hat{e}(i) - e(i) \geq 0; \qquad\qquad (\alpha_2(i) \geq 0);$$

$$\text{(P2.3)} \qquad v(i) - \frac{c(i)}{(1-R^i)} - \tilde{v} \geq 0; \qquad\qquad (\alpha_3(i) \geq 0);$$

which generates the Lagrangian

$$\mathcal{L} = B^i(c(i),e(i),v(i),q(i)) - p_w(w(i) - w_i^o)$$
$$+ \alpha_1(i)(w(i) - c(i)) + \alpha_2(i)(\hat{e}(i) - e(i)) + \alpha_3(i)(v(i) - \frac{c(i)}{(1-R^i)} - \tilde{v}).$$ [19]

The first order conditions are

$$\frac{\partial \mathcal{L}}{\partial c(i)} = B_c^i - \alpha_1(i) - \alpha_3(i)/(1-R^i) \leq 0;$$

$$\frac{\partial \mathcal{L}}{\partial e(i)} = B_e^i - \alpha_2(i) \leq 0;$$

$$\frac{\partial \mathcal{L}}{\partial w(i)} = -p_w + \alpha_1(i) = 0;$$

$$\frac{\partial \mathcal{L}}{\partial c(i)} c(i) = \frac{\partial \mathcal{L}}{\partial e(i)} e(i) = \frac{\partial \mathcal{L}}{\partial \alpha_1(i)} \alpha_1(i) = \frac{\partial \mathcal{L}}{\partial \alpha_2(i)} \alpha_2(i) = \frac{\partial \mathcal{L}}{\partial \alpha_3(i)} \alpha_3(i) = 0 .$$

By assumption (A3), the first order conditions lead to the following allocation rule:

---

[19] Constraint (P2.3) states that there must be enough water available in the system to satisfy the consumptive water right. Note that constraints (P2.2), and (P2.3) implicitly assume that there is some penalty high enough to force compliance to the regulation.

$B_c^i - \alpha_3(i)/(1 - R^i) = p_w$, and $B_e^i - \alpha_2(i) = 0$. There are several reasons why this outcome is not optimal. First, if $\alpha_3(i) = 0$, then there are no binding flow constraints and users equate the marginal benefit from consuming water to a single price. Users do not account for flow effects on other users, nor do they consider the impact of water consumption on water quality under this decision rule. Thus the allocation of water implied by this rule could only be optimal for water allocation in a lentic system with no water quality constraints.

Second, the distribution of discharges is determined by $\hat{e}(i)$. Since water consumption is not optimal in the model, it is not possible to set $\hat{e}(i)$ optimally. This is because damages from $\hat{e}(i)$ depend on water flows. Finally the regulator will not be able to control quality optimally through $\hat{e}(i)$. Assume, for argument's sake, that the regulator is able to set allowable discharges so that the IFN constraint on water quality is met at each site given the equilibrium distribution of water licenses. If there is some perturbation in the water market which would lead to a redistribution of licenses between firms (for example technological change in some industries), there is no reason to believe that the IFN constraints would be satisfied after the redistribution without an accompanying adjustment in allowable discharges between firms. This undermines the power of a decentralized water trading system, since in order to maintain water quality, the regulatory body would be forced to evaluate each transaction and adjust the allowable discharges for each firm. This type of monitoring and adjustment is informationally demanding and not politically feasible.

3.1.b Enforcing IFN constraints on water quality.

Assume that instead of imposing site-specific allowable discharges, the regulator imposes the constraint $q(i) + f^i(c(i), e(i), v(i), q(i)) \geq \tilde{q}$, on each user. The objective function for each user remains the same as (P2.a), however, we replace (P2.2) with the constraint (P2.4) $q(i) + f^i(c(i), e(i), v(i), q(i)) - \tilde{q} \geq 0$ and the Lagrangian becomes

$$\mathcal{L} = B^i(c(i), e(i), v(i), q(i)) - p_w(w(i) - w_i^o)$$
$$+ \alpha_1(i)(w(i) - c(i)) + \alpha_2(i)(q(i) + f^i(c(i), e(i), v(i), q(i)) - \tilde{q}) + \alpha_3(i)(v(i) - \frac{c(i)}{(1 - R^i)} - \tilde{v}).$$

The resulting allocation rule is

$B_c^i - p_w + \alpha_2(i) f_c^i - \alpha_3(i)/(1 - R^i) = 0$, and

$B_e^i + \alpha_2(i) f_e^i = 0$.

If $\alpha_2(i) = \lambda_2(i)$, then $B_e^i + \alpha_2(i) f_e^i = 0$ implies that $\mu_2(i) = 0$. That is, the shadow value of the quality state variable would have to be zero for this policy to be optimal. Even if this were the case, solving for $\alpha_2(i)$, we have

$$B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \alpha_3(i)/(1 - R^i) = p_w,$$

which is equivalent to (9) when $p_w = \mu_1(i)$, and $\alpha_3(i) = \lambda_1(i)$. Since p_w is independent of location, $\Delta\mu_1(i) = 0$, is a necessary condition for the above allocation to be optimal. In general $\Delta\mu_1(i) < 0$. Under a single price regime users equate marginal private benefits, whereas the social cost of water consumption is in fact declining downstream due to spatial externalities.

3.1.c Marketable permits for effluent discharges.

Under this policy, water quality is regulated through a tradable permit system for effluent discharges (as opposed to damages). The regulator allocates a total number of discharge licenses $\bar{H} = \sum_{i=1}^{n} h_i^o$. Each user faces the compliance constraint $h(i) - e(i) \geq 0$ which holds with equality as long as prices are positive (Montgomery, 1972). The objective function for each user becomes

$$\text{(P2.b)} \quad \underset{c(i),e(i),w(i),h(i)}{Max} \quad B^i(c(i),e(i),v(i),q(i)) - p_w(w(i) - w_i^o) - p_h(h(i) - h_i^o).$$

Constraint (P2.2) is now replaced with

$$\text{(P2.5)} \quad h(i) - e(i) \geq 0,$$

and the new Lagrangian is

$$\mathcal{L} = B^i(c(i),e(i),v(i),q(i)) - p_w(w(i) - w_i^o) - p_h(h(i) - h_i^o)$$
$$+ \alpha_1(i)(w(i) - c(i)) + \alpha_2(i)(h(i) - e(i)) + \alpha_3(i)(v(i) - \frac{c(i)}{(1 - R^i)} - \bar{v}).$$

The resulting allocation rule is equivalent to the rule in problem *3.1.a* where $\alpha_2(i) = p_h$. Thus we encounter the same problems as in *3.1.a.* The only difference between the two policies is that users can trade licenses for discharges until the marginal benefits of discharges are equalized. In this case the problem of maintaining instream quality is exacerbated by the fact that users now trade licenses for discharges as well as water.

So far we have been vague about how the regulator would control the aggregate level of damages in the water system so that final water quality is at least $\bar{q}$. In problems *3.1.a* and *3.1.c*, the regulator must be able to calculate the consumption of water and discharges for each user in order to set either $\hat{e}(i)$, or $\bar{H}$, so as to maintain final water quality. In problem *3.1.b*, the regulator sets IFN quality $\tilde{q}$. If $\tilde{q} \geq \bar{q}$, then $q(n+1) \geq \bar{q}$. However, if $\tilde{q} < \bar{q}$ the endpoint constraint will be violated. This problem is avoided if the regulator allocates licenses for total pollution damages, the policy we consider next.

3.1.d Marketable permits for pollution damages.

Assume the regulator allocates licenses for pollution damages $\bar{D} = \sum_{i=1}^{n} d_i^o$ where $\bar{D} = q_o - \bar{q}$. If user i holds licenses to pollute, then $d(i) \geq -f^i(c(i), e(i), v(i), q(i))$.[20] The regulator continues to require that instream flows are maintained at each site.The user's objective function becomes

$$\text{(P2.c)} \quad \underset{c(i),e(i),w(i),d(i)}{Max} \quad B^i(c(i),e(i),v(i),q(i)) - p_w(w(i) - w_i^o) - p_d(d(i) - d_i^o)$$

subject to

$$\text{(P2.6)} \qquad w(i) - c(i) \geq 0; \qquad (\alpha_1(i) \geq 0);$$

$$\text{(P2.7)} \qquad d(i) + f^i(c(i), e(i), v(i), q(i)) \geq 0, \qquad (\alpha_2(i) \geq 0);$$

$$\text{(P2.8)} \qquad v(i) - \frac{c(i)}{(1 - R^i)} - \tilde{v} \geq 0; \qquad (\alpha_3(i) \geq 0);$$

$$\text{(P2.9)} \qquad q(i) + f^i(c(i), e(i), v(i), q(i)) - \tilde{q} \geq 0; \qquad (\alpha_4(i) \geq 0).$$

The Lagrangian is given by

$$\mathcal{L} = B^i(c(i), e(i), v(i), q(i)) - p_w(w(i) - w_i^o) - p_d(d(i) - d_i^o)$$

$$+ \alpha_1(i)(w(i) - c(i)) + \alpha_2(i)(d(i) + f^i(c(i), e(i), v(i), q(i))) + \alpha_3(i)(v(i) - \frac{c(i)}{(1 - R^i)} - \tilde{v})$$

---

[20] Note that damages occur when water quality decreases between users, or when $f^i(c(i), e(i), v(i), q(i)) < 0$.

$$+ \alpha_4(i)\,(q(i) + f^i(c(i),e(i),v(i),q(i)) - \bar{q}).$$

The first order conditions reduce to

$$B_c^i - p_w - \frac{B_e^i}{f_e^i} f_c^i - \alpha_3(i)/(1-R^i) = 0; \text{ and,}$$

$$B_e^i + p_d f_e^i + \alpha_4(i) f_e^i = 0.$$

Note that these conditions are equivalent to (9) and (10) with $p_w = \mu_1(i)$, $p_d = \mu_2(i)$, $\alpha_3(i) = \lambda_1(i)$, and $\alpha_4(i) = \lambda_2(i)$. Since all users face the same prices, p_w and p_d, this policy is optimal only if $\Delta\mu(i)_1 = \Delta\mu(i)_2 = 0$. Therefore it is not optimal in general since the shadow cost of consuming inputs is decreasing downstream.

We have demonstrated that if users are myopic, markets for tradable water licenses are inefficient because users do not internalize the costs of their input decisions on downstream water flows and quality. This decreases value of inputs to downstream users. In the next section we relax the assumption that users are myopic and consider the efficiency properties of a water license market equilibrium when users incorporate the impacts of their own consumption decisions on the spatial pattern of water consumption and pollution in the river system.

3.2 Water License Market Equilibrium with Non-Myopic Users

User i is assumed to recognize that $v(i)$ and $q(i)$ depend on consumption and discharges of users upstream, $(j = 1, \ldots, i-1)$, and takes this into account in deciding whether to purchase from or sell licenses to upstream users. Because $v(i)$ and $q(i)$ do not depend on choices of users downstream from i, an asymmetry is introduced with respect to the willingness to pay for upstream versus downstream licenses. Users distinguish the value of licenses based on the location from which they are purchased. Since $\mu_1(i)$ and $\mu_2(i)$ decrease with i, any market which supports the optimal solution must support separate prices at each location. Therefore, we need to show that in equilibrium there is no incentive for arbitrage, even though the licenses themselves confer identical use rights. In the following section we prove that markets for water licenses and pollution damages operating in conjunction with the enforcement of IFN constraints lead to the optimal allocation of resources if users are not myopic.

Assume that the regulator distributes a fixed total number of water and pollution licenses so that $\bar{W} = v_o - \bar{v}$, and $\bar{D} = q_o - \bar{q}$, respectively. Let w_i^k, and d_i^k, be the numbers of licenses purchased by user i from user k. Negative purchases are interpreted as sales from user i to user k. The final number of holdings for each user is defined by $w(i) \equiv w_i^o + \sum_{k=1}^{n} w_i^k$, and $d(i) \equiv d_i^o + \sum_{k=1}^{n} d_i^k$. Because there is a fixed number of licenses in the system, each user realises that *ceteris paribus*, purchasing a license from user k results in a one unit reduction of k's license holdings. Therefore $\Delta w_i^k = -\Delta w(k)$, and $\Delta d_i^k = -\Delta d(k)$. The number of licenses purchased from user k depends on p_{wi}^k, and p_{di}^k, the prices paid by user i for water and damage licenses respectively from user k. When $k = i$, p_{wi}^k and p_{di}^k are interpreted as the reservation prices at which user i is willing to sell licenses. These are just the shadow costs of holding licenses at site i.

3.2.a Non-Cooperative Nash Equilibrium

The market described above is a pure-exchange Walrasian market. We assume that the license system is in equilibrium when there is no incentive for any pair of users to trade. In this model, each user bids for licenses by offering a price at which he/she is willing to pay for a license from a given location. At the same time, each user identifies a reservation price at which she will sell licenses. Equilibrium is characterized as the Nash outcome to a non-cooperative price game with complete information.[21]

The player set consists of the N users in the water system. The objective of each user is to choose water and damage licenses to maximize user benefits subject to compliance and IFN constraints. Therefore each user chooses vectors $w = (w_i^1, ..., w_i^{i-1}, w_i^{i+1}, ..., w_i^n)$, and $d = (d_i^1, ..., d_i^{i-1}, d_i^{i+1}, ..., d_i^n)$, to solve the programming problem:

$$\text{(P3)} \qquad \underset{c,e,v,q,w,d}{Max} \quad B^i(c(i),e(i),v(i),q(i)) - \sum_{k=1}^{n} p_{wi}^k w_i^k + p_{wi}^i w_i^o - \sum_{k=1}^{n} p_{di}^k d_i^k + p_{di}^i d_i^o$$

subject to the compliance constraints;

$$\text{(P3.1)} \qquad \sum_{k \neq i} w_i^k + w_i^o - c(i) = 0; \qquad (\gamma_1(i) \geq 0);$$

$$\text{(P3.2)} \qquad \sum_{k \neq i} d_i^k + d_i^o + f^i(c(i),e(i),v(i),q(i)) = 0; \qquad (\gamma_2(i) \geq 0);$$

the flow and quality constraints;

$$\text{(P3.3)} \qquad v_o - \sum_{k=1}^{i-1} w(k) - v(i) = 0; \qquad (\gamma_3(i) \geq 0);$$

[21] That is, each player knows the available strategies and payoffs of all other players.

(P3.4) $$q_o - \sum_{k=1}^{i-1} d(k) - q(i) = 0; \qquad (\gamma_4(i) \geq 0);$$

and the IFN constraints;

(P3.5) $$v(i) - \frac{c(i)}{(1-R^i)} - \bar{v} \geq 0; \qquad (\gamma_5(i) \geq 0);$$

(P3.6) $$q(i) + f^i(c(i), e(i), v(i), q(i)) - \tilde{q} \geq 0; \qquad (\gamma_6(i) \geq 0).$$

(P3.3) and (P3.4) reflect the constraints imposed by upstream input consumption on water flows and quality at site i. We assume that no user holds excess licenses in equilibrium, so that (P3.1)-(P3.4) hold as equalities. The Lagrangian for (P3) is given by

$$\begin{aligned}
\mathcal{L} = {} & B^i(c(i), e(i), v(i), q(i)) - \sum_{k=1}^{n} p_{wi}^k w_i^k + p_{wi}^i w_i^o - \sum_{k=1}^{n} p_{di}^k d_i^k + p_{di}^i d_i^o \\
& + \gamma_1(i)[\sum_{k \neq i} w_i^k + w_i^o - c(i)] + \gamma_2(i)[\sum_{k \neq i} d_i^k + d_i^o + f^i(c(i), e(i), v(i), q(i))] \\
& + \gamma_3(i)[v_o - \sum_{k=1}^{i-1} w(k) - v(i)] + \gamma_4(i)[q_o - \sum_{k=1}^{i-1} d(k) - q(i)] \\
& + \gamma_5(i)[v(i) - \frac{c(i)}{(1-R^i)} - \bar{v}] + \gamma_6(i)[q(i) + f^i(c(i), e(i), v(i), q(i)) - \tilde{q}].
\end{aligned}$$

Necessary conditions for equilibrium in the permit market are:

(13) $$\frac{\partial \mathcal{L}}{\partial c(i)} = B_c^i - \gamma_1(i) + \gamma_2(i) f_c^i - \gamma_5(i)/(1-R^i) + \gamma_6(i) f_c^i \leq 0; \qquad \frac{\partial \mathcal{L}}{\partial c(i)} c(i) = 0$$

(14) $$\frac{\partial \mathcal{L}}{\partial e(i)} = B_e^i + \gamma_2(i) f_e^i + \gamma_6(i) f_e^i \leq 0; \qquad \frac{\partial \mathcal{L}}{\partial e(i)} e(i) = 0$$

(15) $$\frac{\partial \mathcal{L}}{\partial v(i)} = B_v^i + \gamma_2(i) f_v^i - \gamma_3(i) + \gamma_5(i) + \gamma_6(i) f_v^i \le 0; \qquad \frac{\partial \mathcal{L}}{\partial v(i)} v(i) = 0$$

(16) $$\frac{\partial \mathcal{L}}{\partial q(i)} = B_q^i + \gamma_2(i) f_q^i - \gamma_4(i) + \gamma_6(i)[1 + f_q^i] \le 0; \qquad \frac{\partial \mathcal{L}}{\partial q(i)} q(i) = 0$$

(17) $$\frac{\partial \mathcal{L}}{\partial w_i^k} = -p_{wi}^k + \gamma_1(i) - \gamma_3(i) \frac{\partial w(k)}{\partial w_i^k} \le 0; \qquad k = 1, \ldots i-1; \qquad \frac{\partial \mathcal{L}}{\partial w_i^k} w_i^k = 0$$

$$= -p_{wi}^k + \gamma_1(i) \le 0; \qquad k = i+1, \ldots n; \qquad \frac{\partial \mathcal{L}}{\partial w_i^k} w_i^k = 0$$

(18) $$\frac{\partial \mathcal{L}}{\partial d_i^k} = -p_{di}^k + \gamma_2(i) - \gamma_4(i) \frac{\partial d(k)}{\partial d_i^k} \le 0; \qquad k = 1, \ldots i-1; \qquad \frac{\partial \mathcal{L}}{\partial d_i^k} d_i^k = 0$$

$$= -p_{di}^k + \gamma_2(i) \le 0; \qquad k = i+1, \ldots n; \qquad \frac{\partial \mathcal{L}}{\partial d_i^k} d_i^k = 0$$

(19) $$\frac{\partial \mathcal{L}}{\partial \gamma_1(i)} = \frac{\partial \mathcal{L}}{\partial \gamma_2(i)} = \frac{\partial \mathcal{L}}{\partial \gamma_3(i)} = \frac{\partial \mathcal{L}}{\partial \gamma_4(i)} = 0;$$

(20) $$\frac{\partial \mathcal{L}}{\partial \gamma_5(i)} \ge 0; \quad \frac{\partial \mathcal{L}}{\partial \gamma_5(i)} \gamma_5(i) = 0; \quad \frac{\partial \mathcal{L}}{\partial \gamma_6(i)} \ge 0; \quad \frac{\partial \mathcal{L}}{\partial \gamma_6(i)} \gamma_6(i) = 0.$$

Given (A1)-(A3), there is an interior solution to each firm's maximization problem. The asymmetry in benefits from purchasing licenses from upstream rather than downstream users is reflected in conditions (17) and (18). Purchasing licenses from any user $k<i$, reduces upstream water consumption and pollution, the value of which is reflected in the multipliers $\gamma_3(i)$, and $\gamma_4(i)$.

The Kuhn Tucker conditions allow us to derive payoff functions for buying permits from any site *k*. The costs of the compliance constraints faced by each firm are given by $\gamma_1(i)$ and $\gamma_2(i)$. (A3) together with (14) implies

$$(21)\qquad \frac{B_e^i}{f_e^i} + \gamma_2(i) + \gamma_6(i) = 0.$$

Using (21) to solve for $\gamma_2(i)$, and substituting this value into (13) yields

$$(22)\qquad B_c^i - \gamma_1(i) - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1-R^i} = 0,$$

which together with (21) determines $\gamma_1(i)$ and $\gamma_2(i)$ given $\gamma_5(i)$ and $\gamma_6(i)$. We can now determine the value of purchasing licenses from site *k*.

If user *i* buys a water (resp. discharge) license from site *k*, then $w_i^k > 0$ (resp. $d_i^k > 0$).[22]

Let $k<i$. Then from (17), $\gamma_1(i) = p_{wi}^k + \gamma_3(i)\frac{\partial w(k)}{\partial w_i^k}$. Substitute the value of $\gamma_2(i)$ derived

---

[22] Note that $w(i)>0$ and $d(i)>0$ does not imply that all of the inequality constraints in (17) and (18) are binding.

from (14) into (15) to generate $\gamma_3(i) = B_v^i - \frac{B_e^i}{f_e^i} f_v^i + \gamma_5(i)$. Making the appropriate substitution into (21) yields

$$(23) \qquad B_c^i + B_v^i - \frac{B_e^i}{f_e^i}[f_c^i + f_v^i] - \frac{\gamma_5(i) R^i}{1 - R^i} - p_{wi}^k = 0.$$ [23]

Equation (23) describes the incentive to purchase upstream water rights, which is the familiar rule equating the marginal benefit of the right to its marginal cost, the price that user i must pay for each license from site k. Similarly, we can combine (14), (16), and (18) to solve for the number of damage licenses which will be purchased by user i from site k:

$$(24) \qquad B_q^i - \frac{B_e^i}{f_e^i}[1 + f_q^i] - p_{di}^k = 0.$$

If $k \geq i$ then (17) and (18) imply $p_{wi}^k = \gamma_1(i)$ and $p_{di}^k = \gamma_2(i)$. Substitution into (21) and (22) yields demand functions for licenses from sites $k \geq i$:

$$(25) \qquad B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1 - R^i} - p_{wi}^k = 0, \text{ and}$$

$$(26) \qquad -\frac{B_e^i}{f_e^i} - \gamma_6(i) - p_{di}^k = 0.$$

Equations (22)-(26) can be used to generate the payoffs to user i from bidding prices p_{wi}^k, or p_{di}^k for licenses from site k.

Definition: *A Strategy for user i $\sigma_i \in \Omega_i$ is a vector of bid prices $\sigma_i = [p_{wi}^1, ..., p_{wi}^k ... p_{wi}^n; p_{di}^1 ... p_{di}^k ... p_{di}^n]$. The set of strategies available to each user, Ω_i,*

---

[23] Recall $\frac{\partial w(k)}{\partial w_i^k} = \frac{\partial d(k)}{\partial d_i^k} = -1.$

is defined on $\Re^{2n}$.

Now consider the following thought experiment. Suppose two users, *i* and *j*, are bidding for licenses from a third user *k*. We show that this bidding game results in the optimal allocation of inputs between users *i*, *j*, and *k*. Since the result holds for any three users *i*, *j*, and *k*, it generalizes to all users in the system. We will adopt the convention that for any $k \neq i$, user *i* wins the game for user *k*'s water licenses if $p_{wi}^{k} > p_{wj}^{k}$, and damage licenses if $p_{di}^{k} > p_{dj}^{k}$. On the other hand, if $k = i$, then user *i* keeps her licenses as long as $p_{wi}^{k} \geq p_{wj}^{k}$, and $p_{di}^{k} \geq p_{dj}^{k}$. Under this convention, each user assumes that in the event of a tie bid, the licenses will go to the other bidder, except when $k = i$, in which case the seller wins the tie.

The payoffs to this bidding game are set out in (27)-(29) below, and are derived from the benefit maximizing conditions given by equations (23)-(26). The payoff from winning a license from user *k* is the marginal benefit of the license, minus the marginal cost which is the price that user *i* pays *k*. The payoff from losing a bid against *j* depends on the location of *j*, since any transfer between *j* and *k*, where *i* is located between the two parties, will result in third party effects due to changes in the return flow regime and pollution loadings up to site *i*.[24] Therefore, the payoffs from winning and losing bids are unequal.

The payoff function has three distinct parts based on the location of user *k* relative

---

[24] We are assuming that second order effects due to differences in natural regeneration of the stream are negligible.

to user i.

A: $\quad$ **k < i**

(27.1) $B_c^i + B_v^i - \frac{B_e^i}{f_e^i}[f_c^i + f_v^i] - \frac{\gamma_5(i) R^i}{1 - R^i} - p_{wi}^k,$ $\quad$ *if* $p_{wi}^k > p_{wj}^k \quad \forall \quad j \neq i;$

(27.2) $B_q^i - \frac{B_e^i}{f_e^i}[1 + f_q^i] - p_{di}^k$ $\quad$ *if* $p_{di}^k > p_{dj}^k \quad \forall \quad j \neq i;$

(27.3) $= \quad 0;$ $\quad$ *if* $p_{wi}^k \leq p_{wj}^k \quad \forall \quad j < i;$

(27.4) $= \quad 0;$ $\quad$ *if* $p_{di}^k \leq p_{dj}^k \quad \forall \quad j < i;$

(27.5) $= \quad B_v^i - \frac{B_e^i}{f_e^i} f_v^i + \gamma_5(i);$ $\quad$ *if* $p_{wi}^k \leq p_{wj}^k \quad \forall \quad j > i;$

(27.6) $= \quad B_q^i - \frac{B_e^i}{f_e^i} f_q^i + \gamma_6(i);$ $\quad$ *if* $p_{di}^k \leq p_{dj}^k \quad \forall \quad j > i;$

Part A describes the payoff from bidding against j for licenses from some user k located upstream from i. The benefit of obtaining upstream water (resp. damage) licenses is equal to the marginal benefit of increased consumption of water (resp. polluting inputs), plus the marginal benefit from the decrease in upstream damages and increased water flows (resp. quality). The benefits from winning the bid are independent of the location of j. If i loses a bid to an upstream user j, then there is no impact on i since this is just a redistribution of water use upstream. On the other hand, if i loses a bid for upstream water (resp. quality) licenses to a downstream user $j > i$, this will lead to a net increase in quality and water flows (resp. quality) at site i even though i does not enjoy the consumptive use of the license.

B: $\mathbf{k=i}$

(28.1) $= p_{wj}^{i} - B_{c}^{i} - B_{v}^{i} + \frac{B_{e}^{i}}{f_{e}^{i}}[f_{c}^{i} + f_{v}^{i}] + \frac{\gamma_5(i) R^{i}}{1-R^{i}};$ $\quad$ *if* $p_{wi}^{k} \le p_{wj}^{k} \quad \forall \quad j<i;$

(28.2) $= -p_{wj}^{i} + B_{c}^{i} + B_{v}^{i} - \frac{B_{e}^{i}}{f_{e}^{i}}[f_{c}^{i} + f_{v}^{i}] - \frac{\gamma_5(i) R^{i}}{1-R^{i}};$ $\quad$ *if* $p_{wi}^{k} > p_{wj}^{k} \quad \forall \quad j<i;$

(28.3) $= p_{dj}^{i} - B_{q}^{i} + \frac{B_{e}^{i}}{f_{e}^{i}}[1 + f_{q}^{i}];$ $\quad$ *if* $p_{di}^{k} \le p_{dj}^{k} \quad \forall \quad j<i;$

(28.4) $= -p_{dj}^{i} + B_{q}^{i} - \frac{B_{e}^{i}}{f_{e}^{i}}[1 + f_{q}^{i}];$ $\quad$ *if* $p_{di}^{k} > p_{dj}^{k} \quad \forall \quad j<i;$

(28.5) $= p_{wj}^{i} - B_{c}^{i} + \frac{B_{e}^{i}}{f_{e}^{i}} f_{c}^{i} + \frac{\gamma_5(i)}{1-R^{i}};$ $\quad$ *if* $p_{wi}^{k} \le p_{wj}^{k} \quad \forall \quad j>i;$

(28.6) $= -p_{wj}^{i} + B_{c}^{i} - \frac{B_{e}^{i}}{f_{e}^{i}} f_{c}^{i} - \frac{\gamma_5(i)}{1-R^{i}};$ $\quad$ *if* $p_{wi}^{k} > p_{wj}^{k} \quad \forall \quad j>i;$

(28.7) $= p_{dj}^{i} + \frac{B_{e}^{i}}{f_{e}^{i}} + \gamma_6(i);$ $\quad$ *if* $p_{di}^{k} \le p_{dj}^{k} \quad \forall \quad j>i;$

(28.8) $= -p_{dj}^{i} - \frac{B_{e}^{i}}{f_{e}^{i}} - \gamma_6(i);$ $\quad$ *if* $p_{di}^{k} > p_{dj}^{k} \quad \forall \quad j>i;$

In Part B, $k=i$, and (28.1)-(28.8) describe the payoffs to i from selling licenses to j. One way of thinking about p_{wi}^{i}, is that it determines the choice of whether or not to sell a water license given some offer p_{wj}^{i}. The same is true for p_{di}^{i}. The payoff functions (28.1)-(28.8) reflect the fact that p_{wj}^{i} and p_{dj}^{i} are the opportunity costs of holding water and damage licenses. Again the payoffs from selling licenses are asymmetric depending on whether user i sells upstream or downstream.

C: **k > i**

$$(29.1) = B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1 - R^i} - p_{wi}^k; \qquad \textit{if } p_{wi}^k > p_{wj}^k \quad \forall \ j \neq i;$$

$$(29.2) = -\frac{B_e^i}{f_e^i} - \gamma_6(i) - p_{di}^k; \qquad \textit{if } p_{di}^k > p_{dj}^k \quad \forall \ j \neq i;$$

$$(29.3) = -B_v^i + \frac{B_e^i}{f_e^i} f_v^i - \gamma_5(i); \qquad \textit{if } p_{wi}^k \leq p_{wj}^k \quad \forall \ j < i;$$

$$(29.4) = -B_q^i + \frac{B_e^i}{f_e^i} f_q^i - \gamma_6(i); \qquad \textit{if } p_{di}^k \leq p_{dj}^k \quad \forall \ j < i;$$

$$(29.5) = 0; \qquad \textit{if } p_{wi}^k \leq p_{wj}^k \quad \forall \ j > i;$$

$$(29.6) = 0; \qquad \textit{if } p_{di}^k \leq p_{dj}^k \quad \forall \ j > i.$$

In Part C, *k* lies downstream from *i*. The construction of the payoffs is analogous to Part A. Note that in Part C user *i* does not benefit from increased flows (resp. quality) if water (resp. damage) licenses are purchased from a downstream site *k*. Nonetheless, *i* will lose flows (resp. quality) that previously went to *k*, if the water (resp. damage) licenses go to some $j<i$.

Given the above payoff function, we can derive the optimal strategy for each user and characterize the Nash equilibrium which results from trading in the permit market.

Definition: *The strategy* $\sigma_i^*(\sigma_{-i})$ *is an* ***Optimal Strategy*** *for user i if for all* $\sigma_i \in \Omega_i$, $\Pi^i(\sigma_i^*, \sigma_{-i}) \geq \Pi^i(\sigma_i, \sigma_{-i})$.

The optimal strategy for user i is derived as follows.

1. (p_{wi}^{i}, p_{di}^{i}) (The Decision to Sell)

From Part B of the payoff function, we see that user i is indifferent between selling water (resp. damage) licenses to upstream users when the price offered is equal to the opportunity cost of holding the license. These conditions define optimal reservation prices for selling water and pollution licenses upstream, $p_{wi}^{i-} = B_c^i + B_v^i - \frac{B_e^i}{f_e^i}[f_c^i + f_v^i] - \frac{\gamma_5(i)R^i}{1-R^i}$, and $p_{di}^{i-} = B_q^i - \frac{B_e^i}{f_e^i}[1 + f_q^i]$ respectively. Similarly, the optimal reservation prices for selling licenses downstream are: $p_{wi}^{i-} = B_c^i - \frac{B_e^i}{f_e^i}f_c^i - \frac{\gamma_5(i)}{1-R^i}$, and $p_{di}^{i-} = -\frac{B_e^i}{f_e^i} - \gamma_6(i)$.

2. (p_{wi}^{k}, p_{di}^{k}) (The Decision to Buy)

Let $p_{wi}^{k\max}(j)$, and $p_{di}^{k\max}(j)$, be the prices at which user i is just indifferent to purchasing the licenses or losing the game to bidder j. If $k<i$, and $j<i$, then from (27.1) and (27.3), we see that the optimal strategy for i in bidding against j for water licenses is to bid $p_{wi}^{k} = p_{wj}^{k} + \varepsilon$, where ε is arbitrarily small, provided $p_{wj}^{k} < p_{wi}^{kmax}(j) = B_c^i + B_v^i - \frac{B_e^i}{f_e^i}[f_c^i + f_v^i] - \frac{\gamma_5(i)R^i}{1-R^i}$. From (27.2) and (27.4), i should bid $p_{di}^{k} = p_{dj}^{k} + \varepsilon$ for a damage license, provided $p_{dj}^{k} < p_{di}^{kmax}(j) = B_q^i - \frac{B_e^i}{f_e^i}[1 + f_q^i]$. If $k<i$ and $j>i$, then from (27.1) and (27.5), it is optimal for i to bid $p_{wi}^{k} = p_{wj}^{k} + \varepsilon$, provided $p_{wj}^{k} < p_{wi}^{kmax}(j) = B_c^i - \frac{B_e^i}{f_e^i}f_c^i - \frac{\gamma_5(i)}{1-R^i}$, and, from (27.2) and (27.6), $p_{di}^{k} = p_{dj}^{k} + \varepsilon$, provided $p_{dj}^{k} < p_{di}^{kmax}(j) = -\frac{B_e^i}{f_e^i} - \gamma_6(i)$. Similarly, if $k > i$, and

$j<i$, then from (29.1) and (29.3) it is optimal to bid $p_{wi}^{k} = p_{wj}^{k} + \varepsilon$, provided $p_{wj}^{k} < p_{wi}^{kmax}(j) = B_c^i + B_v^i - \frac{B_e^i}{f_e^i}[f_c^i + f_v^i] - \frac{\gamma_5(i) R^i}{1 - R^i}$, and, from (29.2) and (29.4), $p_{di}^{k} = p_{dj}^{k} + \varepsilon$, provided $p_{dj}^{k} < p_{di}^{kmax}(j) = B_q^i - \frac{B_e^i}{f_e^i}[1 + f_q^i]$. Finally, if $k>i$ and $j>i$ then from (29.1), (29.5), (29.2), and (29.6), i should bid until $p_{wj}^{k} < p_{wi}^{kmax}(j) = B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1 - R^i}$, and $p_{dj}^{k} < p_{di}^{kmax}(j) = -\frac{B_e^i}{f_e^i} - \gamma_6(i)$.

We now can show, via Lemma 1 and Proposition 1, that a Nash Equilibrium in the trade of water and damage licenses is socially optimal.

Definition: *A set of strategies* $\{\sigma_i^*\}_{i=1}^n$ *is a* ***Pure Strategy Nash Equilibrium*** *if, for all i and* $\sigma_i \in \Omega_i$, $\Pi^i(\sigma_i^*, \sigma_{-i}^*) \geq \Pi^i(\sigma_i, \sigma_{-i}^*)$.

Lemma 1: *A set of strategies* $\{\sigma_i^*\}_{i=1}^n$ *is a* ***Pure Strategy Nash Equilibrium iff***

$$(30) \quad B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1 - R^i} = B_c^{i+1} + B_v^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}}[f_c^{i+1} + f_v^{i+1}] - \frac{\gamma_5(i+1) R^{i+1}}{1 - R^{i+1}}; \text{ and}$$

$$(31) \quad -\frac{B_e^i}{f_e^i} - \gamma_6(i) = B_q^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}}[1 + f_q^{i+1}].$$

Proof: Assume the contrary. Let

$$B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1 - R^i} < B_c^{i+1} + B_v^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}}[f_c^{i+1} + f_v^{i+1}] - \frac{\gamma_5(i+1) R^{i+1}}{1 - R^{i+1}}.$$

This implies that user i's reservation price for selling a water license downstream, given by, $p_{wi}^{i*} = B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1-R^i}$, is less than the marginal benefit to $i+1$ from purchasing the license from i, $B_c^{i+1} + B_v^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}}[f_c^{i+1} + f_v^{i+1}] - \frac{\gamma_5(i+1)R^{i+1}}{1-R^{i+1}} = p_{wi+1}^{i\max}(i)$. Therefore it is optimal for $i+1$ to bid some price $p_{wi+1}^{i} \in (p_{wi}^{i*},\ p_{wi+1}^{i\max}(i)]$, and for i to sell. By the same argument, if $B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1-R^i} > B_c^{i+1} + B_v^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}}[f_c^{i+1} + f_v^{i+1}] - \frac{\gamma_5(i+1)R^{i+1}}{1-R^{i+1}}$ then it is optimal for i to bid some price $p_{wi}^{i+1} \in (p_{wi+1}^{i*},\ p_{wi}^{i+1\max}(i)]$, and for $i+1$ to sell. Only when (30) holds with equality is there no price at which adjacent users will trade to mutual benefit. Therefore (30) is a necessary condition for Nash Equilibrium. By analogy, (31) also is a necessary condition for Nash equilibrium.

The necessary conditions must hold for any adjacent pair of users. In order to show that these conditions are sufficient for equilibrium, we must now show that there is no incentive for non-adjacent trades when the necessary conditions hold. This part of the proof proceeds in two parts.

First, $B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1-R^i} = B_c^{i+1} + B_v^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}}[f_c^{i+1} + f_v^{i+1}] - \frac{\gamma_5(i+1)R^{i+1}}{1-R^{i+1}}$ for all i, implies

(32) $$B_c^{i-1} - \frac{B_e^{i-1}}{f_e^{i-1}} f_c^{i-1} - \frac{\gamma_5(i-1)}{1 - R^{i-1}} > B_c^{i-1} + B_v^{i-1} - \frac{B_e^{i-1}}{f_e^{i-1}} [f_c^{i-1} + f_v^{i-1}] - \frac{\gamma_5(i+1) R^{i-1}}{1 - R^{i-1}} .$$

The LHS of (32), is $p_{wi}^{i-1\,*}$ while the RHS is $p_{wi-1}^{i-1\max}(j)$, when $j>i+1$ which strictly exceeds $p_{wi-1}^{i-1\max}(j)$, when $j<i+1$. Therefore downstream users are never willing to pay the reservation price for non-adjacent upstream water licenses.

Secondly, we show that there is no incentive for upstream users to purchase water licenses from non-adjacent downstream users, even though (30) implies $B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1 - R^i} > p_{wk}^{k-}$, where $k>i+1$. This argument hinges on the fact that if (30) holds, user $i+1$ has an incentive to raise any profitable bid by i to obtain licenses downstream from $i+1$. From the optimal strategies of i and $i+1$, (30) implies $p_{wi}^{k\max}(i+1) = p_{wi+1}^{k\max}(i)$. Therefore there is no price which i is willing to offer $k > i+1$ that will not be raised by $i+1$. Since this is true for any pair i and $i+1$, there is no incentive for non-adjacent trades in water licenses if (30) holds. By analogy there is no incentive for non-adjacent trades in damage licenses when (31) holds. *Q.E.D.*

Proposition 1: *If* $\bar{W} = v_o - \bar{v}$, and $\bar{D} = q_o - \bar{q}$, then the *Nash Equilibrium outcome in the market for water and pollution damage licenses is equivalent to the socially optimal outcome.*

Proof: We prove Proposition 1 in two steps. We first show that Lemma 1 implies (1) - (8) when $p_{wi}^{i*} = \mu_1(i)$, $p_{di}^{i*} = \mu_2(i)$, $\gamma_5(i) = \lambda_1(i)$, and $\gamma_6(i) = \lambda_2(i)$. We then show that

$\bar{W} = v_o - \bar{v}$, and $\bar{D} = q_o - \bar{q}$, imply $p_{wi}^{i*} = \mu_1(i)$, $p_{di}^{i*} = \mu_2(i)$, $\gamma_5(i) = \lambda_1(i)$, and $\gamma_6(i) = \lambda_2(i)$.

Step 1. By Lemma 1, Nash Equilibrium implies (30) and (31). Therefore,

(33) $$B_c^i - \frac{B_e^i}{f_e^i} f_c^i - \frac{\gamma_5(i)}{1 - R^i} - p_{wi}^{i*} = 0; \text{ and}$$

(34) $$-\frac{B_e^i}{f_e^i} - \gamma_6(i) - p_{di}^{i*} = 0.$$ [25]

Equations (33) and (34) imply (1) and (2) when $p_{wi}^{i*} = \mu_1(i)$, $p_{di}^{i*} = \mu_2(i)$, and $\gamma_5(i) = \lambda_1(i)$, $\gamma_6(i) = \lambda_2(i)$. Since (30) and (31) hold for all *i*, we can rewrite the conditions as

(35) $$p_{wi}^{i*} = p_{wi+1}^{(i+1)*} + B_v^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}} f_v^{i+1} + \gamma_5(i+1), \text{ and}$$

(36) $$p_{di}^{i*} = p_{di+1}^{(i+1)*} + B_q^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}} f_q^{i+1} + \gamma_6(i+1),$$

where $p_{wi+1}^{(i+1)*} = B_c^{i+1} - \frac{B_e^{i+1}}{f_e^{i+1}} f_c^{i+1} - \frac{\gamma_5(i+1)}{1 - R^{i+1}}$, and $p_{di+1}^{(i+1)*} = -\frac{B_e^{i+1}}{f_e^{i+1}} - \gamma_6(i+1)$.

Substituting $-\frac{B_e^{i+1}}{f_e^{i+1}} = p_{di+1}^{(i+1)*} + \gamma_6(i+1)$ into (35) and (36) yields

(45) $$p_{wi+1}^{(i+1)*} - p_{wi}^{i*} = -\left[B_v^{i+1} + \gamma_5(i+1) + (p_{di+1}^{(i+1)*} + \gamma_6(i+1)) f_v^{i+1}\right] \text{ and}$$

(46) $$p_{di+1}^{(i+1)*} - p_{di}^{i*} = -\left[B_q^{i+1} + \gamma_6(i+1) + (p_{di+1}^{(i+1)*} + \gamma_6(i+1)) f_q^{i+1}\right]$$

---

[25] Note that the reservation prices p_{wi}^{i*} and p_{di}^{i*} are the shadow values of input consumption at site *i*.

which are equivalent to (3) and (4) when $p_{wi}^{i*} = \mu_1(i)$, $p_{di}^{i*} = \mu_2(i)$, $\gamma_5(i) = \lambda_1(i)$, and $\gamma_6(i) = \lambda_2(i)$.

Note that constraints (P3.5) and (P3.6) are the same as (P1.3) and (P1.4). Therefore (19) implies (5) and (6). Finally (7) and (8) are implied by the definition of $\bar{W}$ and $\bar{D}$.

Step 2. Since the total amount of water and pollution to be allocated is the same under (P1) and (P2), and the necessary and sufficient conditions for Nash Equilibrium imply the first order conditions of (P1), users must consume the same number of inputs in the market equilibrium as in the social optimum. Therefore it must hold that $p_w^i = \mu_1(i)$, $p_d^i = \mu_2(i)$, $\gamma_5(i) = \lambda_1(i)$, and $\gamma_6(i) = \lambda_2(i)$. *QED.*

3.2.b Discussion

Proposition 1 shows that the Nash Equilibrium is socially optimal. The proof follows from the fact that the reservation price at site i for selling water (resp. damage) licenses downstream is equal to the co-state variable (marginal social cost) on water use (resp. emissions) at site *i*. Therefore users choose the socially optimal mix of inputs. The intuition behind this result is as follows. When users are required to hold damage licenses, they become responsible for changes in cumulative water quality and marginal damages from input choices are internalized. Users optimally substitute inputs according to their own user-specific benefit functions and quality conditions at each site. Since users can trade licenses, the external cost of upstream input use is captured in the value of purchasing permits from upstream sites. Therefore users internalize the full social cost of their input decisions, and water and pollutants are allocated to their highest benefit use.

A corollary of this result is that third party effects are an equity rather than an efficiency issue. In our model, third parties are active participants in the bidding process. Bids which reallocate water consumption from downstream to upstream users will be blocked unless the marginal benefits of the trade exceed the marginal costs imposed on third parties.

4. Feasibility

We now address the feasibility of implementing the solution. First we must consider the information requirements for both the regulator and the participants in the market. Secondly, we need to consider the organization of the market, that is, the initial distribution of rights and rules governing exchange.[26] Finally, we must assess the political feasibility of a market in which efficient transactions may require all participants to reorganize their portfolios.

Information requirements for the regulator are high in this model, as incremental damages must be monitored at each site. However, if a water quality production function can be estimated, it will be feasible, although costly, to measure incremental damages. Moreover, similar monitoring costs exist if pollution permits are defined in terms of

---

[26] Ledyard and Szakaly-Moore [14] provide a detailed discussion of alternative trading mechanisms for allocating pollution rights.

emissions, since effluents must be measured at each outflow. Therefore we think that if it is feasible to monitor emissions, it will also be feasible to monitor damages, as long as there are not significant differences in the costs of measuring damages as opposed to discharges.

The transmission of information in the trading regime is another important consideration. The equilibrium described above is a reduced form in the sense that we don't explicitly consider how prices evolve or how information is transmitted in this system. An important feature of the Nash Equilibrium is that players in the market can't be anonymous, that is, each player has to know simultaneously the location and bids of all of the players in the game.

The above considerations suggest that we require a highly organized market. Laissez-faire markets in which the regulator takes a passive role in facilitating trades tend to be inefficient because the transactions costs of finding trading partners are high, and market liquidity is low [14, 11]. In our model a Laissez-faire market has the additional disadvantage of not providing enough information to all of the participants to generate efficient allocations. A trading process which shows promise for addressing the information requirements of the participants in this model is the double auction. In this system, initial allocations are grandfathered to each of the participants who then buy and sell from each other in a central organized market in which all current bids, offers, and trades, are public information. Each participant must outbid the others in order to retain the permits allocated.

While the information requirements on the part of participants in our model are high, the sophistication of current electronic trading systems has the potential to greatly reduce the resulting transactions costs. For example, the Caltech Multiple Unit Double Auction (MUDA) described in Ledyard and Szakaly-Moore [14] offers insight into the potential for electronic trading to facilitate trades in our model.

The MUDA is a highly organized electronic trading system in which participants can continously access and supply bids and offers for multiple units. The advantage of the double auction from an efficiency point of view is that trades occur continuously, and at different prices. This can be contrasted with single price auctions in which bids are sealed and permits are sold at a uniform price. Hahn and Noll [11] argue that sealed bid auctions can reduce monopoly problems which may arise from the initial distribution of permits. However, Ledyard and Szakaly-Moore [14] show that the double auction outperforms single price auctions, even under extreme monopoly conditions. Apart from the information disadvantages, single price sealed auctions are rigid in the sense that they do not allow for any subsequent redistribution of permits. The absence of anonymity in the double auction raises the question of whether players will exercise market power. In uniform price sealed auctions, buyers and sellers have an incentive to misrepresent their bids [21]. However, when units are relatively evenly distributed, the double auction produces close to competitive equilibrium prices and allocations both in theory and in practice [14].

The Nash solution described above assumes that each player takes the actions of the other players as given. In order to provide some justification for this assumption we appeal to the literature on oligopoly behaviour. In this literature, the behaviour of agents must be time consistent in the sense that agents must be able to credibly commit to their equilibrium strategies. The issue of credible commitment in permit markets is important. In our model the total number of permits in the market is fixed, and benefit functions are taken as given. Therefore, users are left to compete in price only. In the absence of any strategic investment which might alter payoffs, there is no way for any user to credibly commit to a non-equilibrium strategy. Thus the non-cooperative game described above can be seen as a variant of the one-shot Bertrand price game which results in a competitive price outcome.

Markets in which the regulatory agency sells permits and retains revenue will make those who should reduce pollution on efficiency grounds worse off. Therefore these organizations are not considered politically viable [14]. In our model, the efficiency problems from third party effects are eliminated because all users have the opportunity to enter the market and adjust their portfolios. However, these effects are costly from a political point of view since trades will still generate either third party gains or losses. A necessary question for future research is whether or not there is some compensation mechanism, perhaps through the initial allocation, which will neutralize these gains and losses and increase the political viability of our trading system.

Finally, the greatest obstacle to implementation of our model is the institutional context in which water is currently allocated. States which allow trades in water do so with many restrictions attached. Protection from third party injury is established under common law. At the same time, definitions of third party damages are expanding to include the public interest (see [1, 6, 8, 16, 25]). In our model public interest is modelled passively through instream flow and quality constraints. It would be useful to directly incorporate public interest in our model.

5. Conclusions

We have derived necessary and sufficient conditions for maximizing the total social benefit of consumptive water use and pollution discharge in a river system. A marketable permit system for surface water rights coupled with a permit system for allocating pollution damages is allocatively efficient if users recognize that there are asymmetric benefits from purchasing upstream versus downstream licenses. An important contribution of our paper is the proof that markets can support location specific prices for environmental rights.

The main conclusion of this analysis is that water control agencies must consider the combined effects of water consumption and discharges when designing institutions for allocating water and maintaining water quality. In commenting on the feasibility of implementing the results from the model, we note that more research is necessary on the institutional setting in which permits will be traded, particularly on how information is transferred and how compensation mechanisms can mitigate third party effects.

6. References

[1] T.L. Anderson and P. Snyder, Water markets: priming the invisible pump", Cato Institute, Washington, D.C., (1997).

[2] C. Blackorby, D. Nissen, D. Primont, and R.R. Russell, Recursively decentralized decision making, *Econometrica* **42**, 487-95 (1974).

[3] J.B. Braden, G.V. Johnson, A.Bouzaher, and D. Miltz, Optimal spatial management of agricultural pollution, *American Journal of Agricultural Economics* **71**, 404-13, (1989).

[4] H.S. Burness and J.P. Quirk, Appropriative water rights and the efficient allocation of resources, *American Economic Review* **69**, 25-37 (1979).

[5] H.S. Burness and J.P. Quirk, Water law, water transfers, and economic efficiency: the Colorado River, *Journal of Law and Economics* **23**, 111-134 (1980).

[6] B.G. Colby, Water reallocation and valuation: voluntary and involuntary transfers in the western United States, in "Water law: trends, policies and practice" (K.M. Carr and J.D. Crammond Eds.), American Bar Association, Section of Natural Resources, Energy and Environmental Law, Chicago Il, (1995).

[7] M.S. Ejaz and R.C. Peralta, Modeling for optimal management of agriculture and domestic waste water loading to streams, *Water Resources Research* **31**, 1087-96 (1995).

[8] G.A. Gould, Recent developments in the transfer of water rights, in "Water law: trends, policies and practice", *op. cit.*, (1995).

[9] L.M. Hartman and D. Seastone, "Water transfers: economic efficiency and alternative institutions", John Hopkins Press, Baltimore, MD, (1970).

[10] R. Hahn, Market power and transferable property rights, *Quarterly Journal of Economics* **99**, 735-65 (1984).

[11] R. Hahn and R. Noll, Barriers to implementing tradable air pollution permits: problems of regulatory interactions, *Yale Journal on Regulation* **1**, 63-91 (1983).

[12] R.N. Johnson, M. Gisser, and M. Werner, The definition of a surface water right and transferability, *Journal of Law and Economics* **24**, 273-88 (1981).

[13] A. Keeler, Noncompliant firms in TDP markets, *Journal of Environmental Economics and Management* **21**,180-89 (1991).

[14] J.O. Ledyard and K. Szakaly-Moore, Designing organizations for trading pollution rights, *Journal of Economic Behavior and Organization* **25**, 167-96 (1994).

[15] D. Leonard and N.V. Long, "Optimal control theory and static optimization in economics", Cambridge University Press, Cambridge, MA., (1992).

[16] M.H. Lennihan, The California drought emergency water bank: a successful institutional response to severe drought, in "Water law: trends, policies and practice" *op. cit*, (1995).

[17] A. Malik, Markets for pollution control when firms are noncompliant, *Journal of Environmental Economics and Management* **18**, 97-106 (1990).

[18] A. McGartland and W. Oates, Marketable permits for the prevention of environmental deterioration, *Journal of Environmental Economics and Management* **12**, 207-228 (1985).

[19] W. Montgomery, Markets in licenses and efficient pollution control programs, *Journal of Economic Theory* **5**, 395-418 (1972).

[20] C.J. Myers and R.A. Posner, Market transfers of water rights: toward an improved market in water resources, National Water Commission, Legal Study No. 4, NTIS No. NWC-L-71-009, (1971).

[21] R. Myerson and M. Satterthwaite, Efficient mechanisms for bilateral trading, *Journal of Economic Theory*, **29**, 107-33 (1983).

[22] Northern River Basins Study, Combined effects of dissolved oxygen level and bleached kraft pulp mill effluent and municipal sewage on a mayfly (*Baetis tricaudata*): assessments using artificial streams, Northern River Basins Study Project Report No. 98, Edmonton, Alberta, (1996).

[23] M. Reisner, "Cadillac desert: the American west and its disappearing water", Douglas and McIntyre, (1993).

[24] N. Spulber and A. Sabbaghi, "Economics of water resources: from regulation to privatization", Kluwer Academic Publishers, Boston MA., (1994).

[25] A. Tarlock, Reallocation: it really is here, in "Water law: trends, policies and practice"*op. cit.*, (1995).

[26] K. Unger, Locational pricing of an environmental input, *Journal of Environmental Economics and Management* **5**, 207-19 (1978).

[27] R.D.C. Wong, and J.W. Eheart, Market simulations for irrigation water rights: a hypothetical case, *Water Resources Research* **19**, 1127-38 (1983).

[28] A.P. Xepapadeas, Observability and choice of instrument mix in the control of externalities, *Journal of Public Economics* **56**, 485-498 (1995).

Concluding Remarks

The three essays presented in this dissertation contribute to our understanding of the efficiency of tradable permit systems when they are introduced in real life settings. Market instruments are increasingly being used to allocate environmental inputs. Both regulators and firms find market systems attractive because the decentralized decision making aspect of these systems maximizes the flexibility of firms in choosing cost minimizing methods of compliance. Market systems also minimize the information requirements of the regulator once the initial regulation has been designed. The essays presented in this dissertation have direct policy implications for the design of marketable permit systems and suggest avenues for future research.

Both market power and noncompliance are factors which can influence the effectiveness of tradable permit systems [3, 5]. In Essay 1, we show that the net social benefits from emission reductions are not independent of the initial allocation of rights to pollute when a firm has market power in the permit market, a result obtained by [2]. However, by incorporating noncompliance in the model we show that in addition to determining aggregate abatement costs, the initial allocation of rights directly affects enforcement costs. The results show that the initial allocation of permits and the parameters of the expected penalty function are complementary instruments which should be chosen simultaneously by the regulator to maximize the social benefits from emission reductions. Further research on this matter would be useful.

In Essay 2, we develop a theoretical model which reveals a tendency for firms to over-invest in abatement capital relative to the least cost solution. In our model, risk aversion on the part of managers making compliance decisions distorts the least cost abatement strategy of a firm. Our model is consistent with stylized facts of the EPA's sulfur dioxide emissions trading program where permit prices are lower than expected, while investment in scrubbers and total compliance costs are higher than necessary [1]. The results from our model show that first-best costs are not feasible when there is moral hazard. Policy makers who do not consider moral hazard will underestimate the aggregate costs of compliance, and overestimate the demand (and equilibrium price) for permits. Thus the overinvestment in scrubbers and the low price of permits observed under Phase I of the EPA's trading program does not reflect a design problem associated with the permit market. Instead it indicates that the regulatory agency should account for moral hazard, both when setting initial emissions targets and when tracking the performance of the market. This research suggests that further empirical investigation into the appropriate parameters and functional forms which describe the utility functions of managers as well as the compliance costs of firms is necessary to estimate accurately aggregate compliance costs and least cost strategies.

In Essay 3, we consider the problem of designing regulations when there are interdependent objectives. We consider the problem of allocating both consumptive water rights and pollution inputs along a river. We derive several "negative" results concerning the ability of a regulator to regulate water quality when water rights are traded. We show that combined marketable permit systems for allocating rights to consume water and to pollute are efficient under stringent assumptions about the behaviour of users along the stream. These assumptions effectively require that users be fully informed about the trading activity of all other users in the system. Equity considerations may undermine the political and legal feasibility of implementing an economically efficient system for allocating these rights. This essay generates interesting questions regarding the design of institutions that could implement the tradable permit system developed in the model. Further research is indicated to assess whether the initial allocation can be used as a tool to increase the political feasibility of the system. In addition, further research into how information is transmitted in the system through the rules governing the trade of permits is necessary.

The main conclusion to be drawn from this dissertation is that small deviations from the "textbook" model of a tradable permit system [4], can result in unanticipated outcomes. Recognizing institutional details about the regulatory and market environment when designing a trading system can greatly improve its performance. Because of the complexity of the real world, there is a need to continue to investigate the properties of these systems in particular applications.

References

[1] D.R. Bohi and D. Burtraw, SO_2 allowance trading: how experience and expectations measure up, Resources for the Future, Discussion Paper 97-24 (1997).

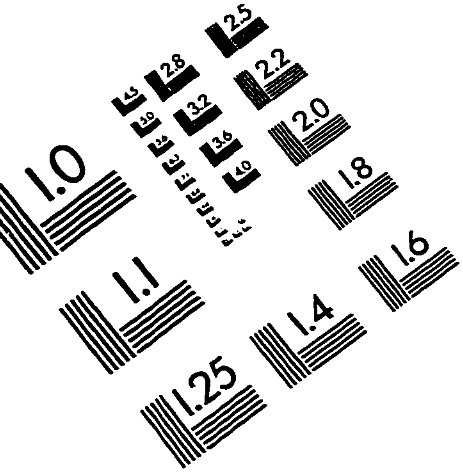
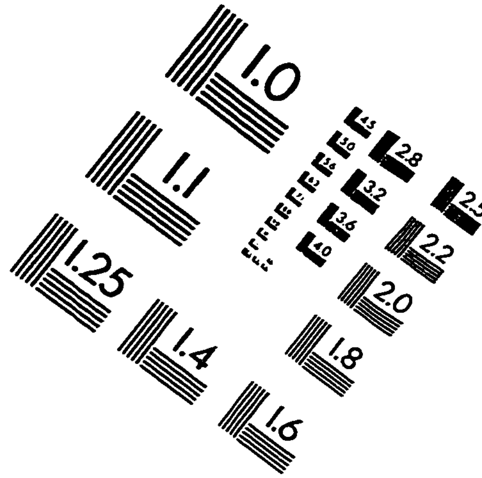
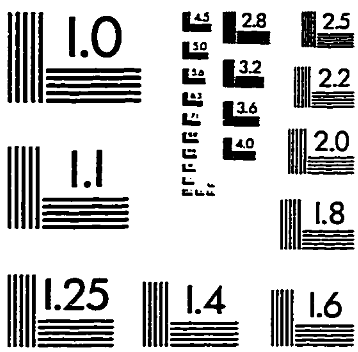
[2] R. Hahn, Market power and transferable property rights, *Quarterly Journal of Economics* **99**, 735-765 (1984).

[3] R. Hahn and R. Noll, Designing a market for tradable emission permits, in "Reform of environmental regulation" (W. Magat, Ed.) Ballinger, Cambridge, MA (1982).

[4] W. Montgomery, Markets in licenses and efficient pollution control programs, *Journal of Economic Theory* **5**, 395-418 (1972).

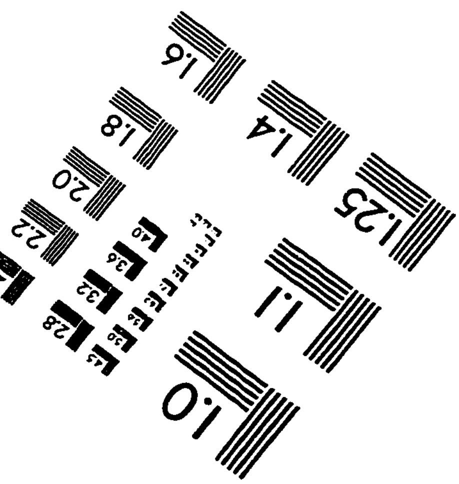
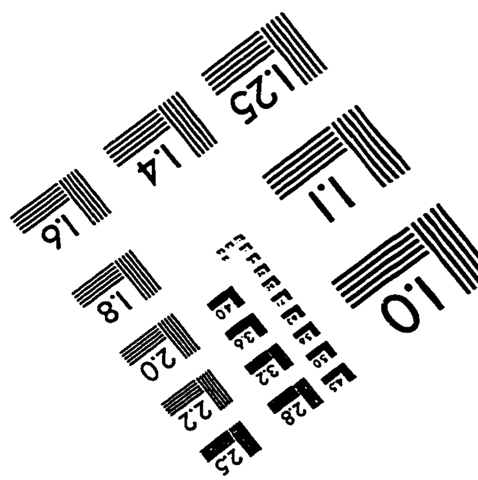
[5] C.S. Russell, W. Harrington and W. J. Vaughan, "Enforcing pollution control laws", Resources for the Future, Washington, DC (1986).

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6″

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989