Dynamic moral hazard and optimal regulation of non point source pollution

Andrea La Nauze* Claudio Mezzetti*

December 5, 2014

Abstract

Non point sources of pollution such as agriculture produce diffuse emissions which are unmeasurable, arise from unobservable actions and cause stochastic pollution of air and waterways. Despite the significant damage they cause, regulation of these sources has proven far from straightforward in both theory and practice. To ensure tractability, theoretical treatments remain either static or focus on determining the optimal level of pollution. In contrast, we specify a rich model of firm incentives in a dynamic setting. Specifically, using the continuous time methods pioneered by Sannikov (2008) we treat (multiple) firms as having limited liability and outside options. We model a regulator who must design an incentive compatible contract to implement abatement and show that the solution is a simple incentive mechanism. The regulator collects a constant tax and links a firms' *promised* payoff to the flow of pollution. A firm is paid a rebate once this promised payoff exceeds their social product. We show that in this contract the regulator can balance its expected budget and the firm receives its social product. We then contrast the contract to a tax on the stock of pollution.

Introduction

Emissions from non point sources pollute air and waterways yet they largely remain unregulated. Non point sources produce diffuse emissions, emissions that cannot be attributed to a source and

*Department of Economics, University of Melbourne, 111 Barry Street, Carlton, Victoria 3053, Australia. The respective email addresses of the authors are: ala1@unimelb.edu.au and cmez@unimelb.edu.au.

result in random observed pollution levels. They arise in a range of sectors including transport, agriculture, mining and forestry and have significant impacts on air and water quality.

The failure to regulate non point sources is due neither to a lack of policy nor academic focus on the problem. The impact of non point source emissions is well recognized. Perhaps the most important impact of diffuse emissions is poor water quality. Diffuse emissions caused by runoff and/or leaching (Fisher-Vanden and Olmstead, 2013) are responsible for polluting waterways. Runoff, for example, transports nutrients (such as phosphorus and nitrogen) and sediments to receiving waterways. These deposits are the primary cause of waterway eutrophication resulting in harmful algal blooms and hypoxia (oxygen depletion) which can destroy aquatic life and threatens ecosystems. Globally, the World Resources Institute suggests that 762 coastal areas are known to have been impacted by eutrophication and/or hypoxia (WRI, 2011). In the United States, pollution resulting from diffuse emissions has been named the greatest threat to the health of rivers and streams (EPA, 2009). Seasonal hypoxia has affected Puget Sound, the Gulf of Mexico and Chesapeake Bay. Despite this, the relevant legislative tool in the United States, the Federal Water Pollution Control Act (the Clean Water Act), does not provide for the regulation of diffuse emissions or non point sources.

Alongside significant policy interest is an extensive academic literature devoted to studying the regulation of non point sources (see Shortle and Horan (2002), Xepapadeas (2011) and Fisher-Vanden and Olmstead (2013) for surveys). Unlike the bulk of this literature we adopt the framework of a principal agent model to characterise the regulation of non point sources as a problem of dynamic moral hazard. In this model the moral hazard problem arises because pollution is uncertain and a polluter's actions are unobservable. We model the accumulation of pollution as a brownian motion whose drift is affected by the actions of polluting firms (the agents). The polluters are risk neutral but suffer from limited liability. In our principal agent model the regulator must design the optimal set of incentives to induce abatement and respect three constraints: the participation constraint, the incentive compatibility constraint and the limited liability constraint.

The principal agent approach to studying environmental problems is not in itself new. Chambers and Quiggin (1996) for example adopt a principal agent framework to study non point source pollution and Laffont (1995) to study environmental risk. Neither considers the dynamics of the principal agent relationship. Xepapadeas (1992) and Athanassoglou (2010), who following Segerson (1988) study taxes on the observed stock of pollution (ambient taxes), both model diffuse emissions as a brownian motion in a dynamic environment. They do not however consider an environment with a limited liability or an individual rationality constraint.

Despite the scale of the problem and the list of well studied instruments, regulation of non point sources is sparse. Ambient taxes for example have been adopted only in a handful of instances.[1] One potential explanation is that ambient taxes can result in large transfers to and from regulated firms hence would expose firms to considerable risk if pollution is volatile (Karp, 2005). We specify a model where the firm has limited liability and so volatility is costly. We also explicitly account for a firm's ability to exercise their outside option. Our model is thus appropriate for studying regulation in environments where firms face credit constraints and have the option to cease operations and move elsewhere to avoid regulation.

In Section 1 we outline our model of a regulator (the principal) who must design an incentive scheme (a contract) to induce polluting firms to undertake abatement. As indicated, the regulator cannot observe the firms' abatement actions, they observe only the ambient stock of pollution (or level of water quality) from which they suffer damage. We consider the case for which it is optimal for the regulator to keep the stock of pollution constant. Importantly, the regulator and the firms contract over an infinite horizon. This dynamic setting is an important feature of the policy environment, mimicking the ongoing interactions between regulated firms and regulatory agencies. The dynamic setting is also important to the structure of the optimal contract: it enables the regulator to motivate the firm with both short and long run incentives.

The firms in our model can be thought of as mines, farms and forestry operations whose runoff affects the water quality of a local river either by increasing nutrient or sediment loads. By its nature this runoff is diffuse, entering the river at many unknown points, and neither the volume of runoff nor its content can be measured at its source. Firms can undertake costly abatement actions, for example reduce fertiliser, use less intensive machinery or construct riparian buffers. It is either too costly to verify these actions (for example inspecting every riparian zone on a property), or they are unverifiable (for example assessing the volume of fertiliser applied to a crop).

The volatility of pollution and hence observed water quality arises due to random weather events. To provide some intuition for this volatility consider the following example: a farmer always applies the same quantity of fertiliser to a crop. If there is no rain then the regulator observes no change in water quality; however if it does rain then a change in water quality (for

[1] Karp (2005) outlines such a handful of examples where ambient-type schemes have been implemented.

example an increase in nitrogen levels) is detected. In this example, the same action results in a different measured level of ambient pollution depending on a random weather event. The regulator is unable to distinguish whether the farmer used excessive fertiliser or whether rain affected the rate at which the fertiliser was delivered to the stream. Similarly, the same excavation practices might result in minimal sediment run off from a mine without rain, however strong rain causes significant erosion and sedimentation of a near-by stream.

In Section 2 we derive the optimal regulatory contract in a setting in which firms have limited liability. We show that this contract combines a constant tax with a dynamic rating of firm performance and a rebate. The dynamic rating ties each firm's payoff to the drift of pollution. Increases in the ambient level of pollution cause the rating to fall. Once the rating exceeds their performance threshold, the firm receives a rebate. Below this threshold a firm pays a constant tax. We provide comparative statics for the regulator's value and demonstrate conditions under which it is optimal to hold the pollution stock constant. In Section 3 we show that when the regulator is forced to balance its expected budget (i.e. be revenue neutral), each firm's performance threshold is equivalent to the social value of a its production (i.e. profits minus the cost of abatement). Under this contract the total expected value of a firm is also equal to the social value of their production. We then show how the tax rate varies with parameters of the model.

In Section 4 we compare the contract to a benchmark setting in which firms do not have limited liability. In this benchmark setting a firm's current (i.e. not their promised payoff) is tied to the flow of pollution. We then contrast our two contracts with an ambient tax. Rather than incentivising the firm through a tax on the stock of pollution as in an ambient tax, both contracts tax the flow of pollution.

To date the application of dynamic contracting models has largely been in corporate finance and has focussed on a single agent (see for example Clementi and Hopenhayn (2006), Albuquerque and Hopenhayn (2004), Biais et al. (2010)). Our paper studies dynamic contracting in a novel setting. To do so we draw on the approach outlined in DeMarzo and Sannikov (2006) and Sannikov (2008) who demonstrate the tractability of models of dynamic moral hazard in continuous time. Our setting differs in a few key respects to those studied in the dynamic moral hazard literature. Firstly, we focus on a setting with multiple agents who may enter and exit over an infinite horizon. Secondly, in our model the limited liability constraint takes a different form. We allow the instantaneous payoff of a firm to be negative but bounded from below. Finally we constrain the principal to ensure that they guarantee the firm their expected social product. Our model also generates several new insights. Unlike existing models where the performance threshold must be solved computationally, we find a closed form solution by imposing budget balance. We show that doing so determines the endogenous tax rate of the firm. The paper is structured as follows. Section 1 outlines the model. Section 2 derives the optimal contract. Section 3 explores the notion of budget balance and computes the tax rate of the firm. Section 4 derives the optimal contract without limited liability and finally Section 5 concludes.

1 The model

In this section we introduce a principal agent model of non point source regulation. The key characteristics of the non point problem are the randomness of the level of pollution and the hidden actions of polluters. Most environmental regulation occurs in a repeated setting and hence it is appropriate to model the interaction between regulator and firms as a dynamic one. Dynamic contracts can differ significantly to those in static settings, yet can be intractable. Continuous time methods provide a convenient and tractable machinery (DeMarzo and Sannikov, 2006; Sannikov, 2008). We thus formulate our model as a dynamic continuous time agency problem.

Consider an environmental regulator (the principal) and a group of regulated firms (the agents). All are risk neutral however firms have limited liability. The firms can be thought of as mines, farms or forestry operations. The activities of firm i generate a profit flow π^i and affect the water quality of a local river either by increasing nutrient or sediment loads. Each firm can undertake a costly abatement action, for example reduce fertiliser use, adopt less intensive tillage practices or construct riparian buffers. In a deterministic world abatement could eliminate pollution. However actual pollution is stochastic and hence firms do not have the ability to completely control it.

We assume that firms can either pollute or abate. If a firm abates they produce zero net emissions. Hence we model firm i at time t as choosing an abatement action $a_t^i \in \{0, 1\}$. The flow cost of abatement ($a_t^i = 1$) is given by c^i. We assume that $\pi^i > c^i$ so that it is still profitable for the firm to produce whilst abating.

The drift of pollution is a function of each firm's abatement choice. We allow firms to enter and exit. Let $\mathbf{N_t}$ be the set of firms that have entered by time t, and N_t by the number of firms that have entered by t. Define $\mathbf{M_t} \subseteq \mathbf{N_t}$ as the set of firms not polluting at time t and let M_t

be the number of firms not polluting at t. Firms who have entered are those who are active and those who have exited. Firms not polluting at time t are firms who have exited, and firms who are abating. We specify the evolution of pollution as:

$$dY_t = \mu_t + \sigma dZ_t \tag{1}$$

where μ_t is the drift of pollution, dZ_t is a standard brownian motion (a shock common to all firms) and σ is the volatility of pollution. The drift μ_t of pollution is given by:

$$\mu_t = [p(\mathbf{N_t}) - p(\mathbf{M_t})] \tag{2}$$

where $p(.)$ is a function such that $p(0) = 0$. Then if N_t firms are not polluting at time t, the drift of pollution $\mu_t = 0$. If only firm i pollutes then the drift of pollution $\mu_t = p(\mathbf{N_t}) - p(\mathbf{N_{-i}})$ where $\mathbf{M_t} = \mathbf{N_{-i}} = \mathbf{N_t}\backslash i$ is the set of all firms except i. Define $A^i = p(\mathbf{N_t}) - p(\mathbf{N_{-i}})$. Thus $\mu_t = A^i$ if only firm i pollutes.

The regulator observes dY_t, but the abatement choices of firms are hidden by the distortion created by the brownian motion. The higher the level of volatility σ, the less information the regulator can deduce from the observed flow dY_t. The distortion represents random weather events such as precipitation that affect the flow of pollution measured by the regulator. This flow is independent of the stock of pollution Y_t. We interpret the shock to pollution as arising from random weather events which are independent of the stock of pollution.[2]

The regulator motivates each firm to abate by the transfer dI_t^i. The transfer can either be from the regulator to a firm ($dI_t^i > 0$) or vice versa from the firm to the regulator ($dI_t^i < 0$). The next section focuses on determining the form of dI_t^i. Given this transfer, the payoff to the firm from their abatement choice a_t^i evolves according to:

$$d\Pi_t^i(a_t^i) = (\pi^i - c^i a_t^i)dt + dI_t^i \tag{3}$$

If the firm chooses to abate then the flow of their payoff is $(\pi^i - c^i)dt + dI_t^i$. In practice firms typically

[2] Xepapadeas (1992) and Athanassoglou (2010) also formulate non point source pollution as a brownian motion however in their models the volatility of pollution is a function of the current stock.

have limited liability; that is, their instantaneous payoff must satisfy the constraint $d\Pi_t^i \geq -l^i dt$ where $l^i dt$ represents some exogenous credit limit of the firm. This places a constraint on the transfer such that $dI_t^i \geq -(l^i + (\pi^i - c^i))dt$. We also assume that there is a policy constraint on the regulator's problem such that $dI_t^i \geq -\tau^i dt$. This represents a constraint on the amount that a regulator can take from a firm in any instant. Thus τ^i can be interpreted as an exogenously given maximum fine or license fee that the regulator can levy. We assume that $-\tau^i \geq -(l^i + (\pi^i - c^i))$ so that the credit limit constraint is not binding. The firm's total expected payoff at time t is the discounted expected value of all future instantaneous payoffs $d\Pi_t^i$. Let W_0^i be the value of this expected future payoff for firm i at time 0. Then

$$W_0^i = E_0 \left[\int_0^{T^i} \exp^{-\gamma^i s} d\Pi_s^i + \exp^{-\gamma^i T^i} R^i \right] \tag{4}$$

where the expectation is taken at $t = 0$ and firm i discounts at rate γ^i. The firm exits at endogenously chosen time T^i when they exercise their outside option R^i (for example transferring business to another state).

At time $t = 0$, W_0^i is firm i's continuation value, its value conditional upon the contract continuing at time $t = 0$. Let W_t^i be i's continuation value, or total expected discounted payoff from time t, conditional on not having exited (i.e. not exercised their outside option) prior to t, then:

$$W_t^i = E_t \left[\exp^{\gamma^i t} \int_t^{T^i} \exp^{-\gamma^i s} d\Pi_s^i(a_s^i) + \exp^{-\gamma^i (T^i - t)} R^i \right] \tag{5}$$

The firm's continuation value depends on its abatement action a_t^i.

We focus on the case where it is optimal to induce all (active) firms to abate at each point in time i.e. $a_t^i = 1$ for all i and $t < T^i$. In turn, this keeps the stock of pollution constant at the initial level Y_0. In our simplified model, the regulator must design a contract to ensure that active firms always abate.[3] This contract maximises the regulator's value which is the discounted expected cost of transfers dI_t^i. These transfers must in turn satisfy the policy constraint $dI_t^i \geq -\tau^i dt$, which is a legislated maximum fee the regulator can impose on firm i.

[3]Later we derive the conditions under which it is optimal to design the contract to always implement abatement.

For now we take τ^i as exogenous, we later go on to determine τ^i by imposing a budget balance condition on the regulator. This budget balance condition will stipulate that the total expected payoff to the regulator (or the present value of the expected net transfers to the regulator from the firm) must be zero. This is a convenient simplification of the true social planner's problem. The social planner would choose the path of abatement and a set of transfers that maximise social welfare, in other words to balance the damage of pollution with the cost of implementing abatement at any point. We simplify the problem in two ways: first, by focusing on the case where abatement is always optimal. Second, by re-writing the objective function as a constrained cost minimisation problem on the part of the regulator. Our simplified approach allows us to focus on the structure of the dynamic incentives rather than on determining the optimal path of abatement.

The total expected value to the regulator at time $t = 0$ is given by:

$$b_0 = E_0 \left[\int_0^{T^i} \cdots \int_0^{T^N} - \exp^{-\delta s} \sum_{i=1}^{N} dIs^i \right] \tag{6}$$

where the regulator discounts at rate $\delta < \gamma^i \; \forall \; i$. The regulator is thus more patient than the most patient firm, reflecting the fact that the regulator cares more about future generations than does any firm. The optimal contract maximises the regulator's value b_0 subject to the constraints $dI^i \geq -\tau^i dt$ and $b_0 = 0$.

2 The contract

The optimal contract specifies a dynamic transfer dI_t^i and recommends the abatement action $a_t^i = 1$ for all t and i. Given the repeated setting, the optimal contract could tie each firm's payoff to the entire history of pollution at time t. This makes the regulator's problem very complex. However Spear and Srivastava (1987) show in a single agent, discrete time setting that the regulator's problem can be reduced to a static variational problem with the continuation value of the firm as a state variable. If the continuation value of a single firm summarises all relevant information on the entire history of their actions, it must do so for each firm in a multi-agent setting. The continuation values W_t^i are therefore state variables of the regulator's problem.

Here, following the approach of DeMarzo and Sannikov (2006) and Sannikov (2008), the state variables W_t^i reduce the problem to one of optimal stochastic control. In other words, W_t^i summarises all relevant information on a firms' actions up to time t and serve as a dynamic rating of past performance. The continuation value W_t^i is firm i's expected future payoff, it can also be interpreted as the future value that the firm is promised by the regulator. At time t the firm has two types of payoffs: its instantaneous payoff $d\Pi_t^i$ and its promised payoff W_t^i. The regulator can use these payoffs as imperfect substitutes to provide each firm with incentives to abate.

The regulator's value function is the solution to the Hamilton-Jacobi-Bellman (HJB) equation of the stochastic control problem. In this problem the HJB has state variables W_t^i. To solve the HJB equation we first need to specify the evolution of W_t^i.

2.1 The evolution of the firm's value

To guarantee that the firm abates, the regulator must ensure that the total benefit that the firm receives from abating is greater than the total benefit they receive from polluting. But the regulator cannot observe a firm's chosen action, it observes only the noisy signal dY_t. The firm's total benefit includes its current payoff and its future payoff. For incentive compatibility, one or both of these payoffs must be linked to the realisation of pollution. The firm's current payoff is constrained by limited liability, but shocks to pollution arise from a brownian motion with unbounded variability. In the optimal contract therefore, the regulator provides incentives to firm i by tying the evolution of its *promised value* (i.e. dW_t^i) to the flow of pollution. Lemma 1 formalises this notion.

Lemma 1. *At time $t < T^i$ there is a sensitivity β_t^i of firm i's continuation value toward their action via the flow of pollution such that:*

$$dW_t^i = \gamma^i W_t^i dt - (\pi^i - \lambda^i A^i)dt - dI_t^i + \beta_t^i dY_t \tag{7}$$

where

$$\beta_t^i \leq -\lambda^i \tag{8}$$

W_t^i follows this process as long as $t < T^i$ where T^i is the first time at which $W_t^i < R^i$.

The derivation of (7) uses the martingale methods introduced in the principal agent literature by Sannikov (2008) and DeMarzo and Sannikov (2006) and is provided in the appendix. The process dW_t^i can be interpreted as the regulator's 'promise keeping constraint'. Recall that at time $t = 0$ the value W_0^i equals the total expected value to firm i. This is the total value that the regulator promises to firm i via the terms of the contract; it must satisfy the firm's individual rationality constraint. That is, $W_0^i \geq R^i$. In expectation if the firm accepts the contract and abates then the regulator must deliver this value to the firm. To keep the firm indifferent between payments now and in the future, the firm's promised value grows over time at their discount rate γ^i. Value received by a firm at time t (i.e. $(\pi^i - c^i)dt + dI_t$) constitutes part of its promised value and reduces the remaining amount that the regulator must guarantee to the firm. The firm exits at T^i, the first time at which $W_T^i < R^i$. That is, the firm exits at time T^i when their promised future value is less than their outside option R^i.

To ensure the firm abates, the promised value is also linked to the flow of pollution dY_t. The parameter β_t^i governs how sensitive firm i's promised value is to the signal dY_t. In other words, it governs how firm i's future benefits change with the observed flow of pollution. The parameter $\lambda^i = \frac{c^i}{A^i}$ is the unit cost of abatement, the total cost divided by the level of abatement. To demonstrate that $\beta_t^i \leq -\lambda^i$ we focus on the firm's incentives to abate. We look for a Nash equilibrium amongst the firms, that is, we consider the firm's incentives to abate conditional on all other active firms choosing to abate. For incentive compatibility it is sufficient to derive conditions under which firm i chooses $a_t^i = 1$ at any arbitrary moment in time.

At time t the firm must balance its payoff from deviating today with the reduction in its payoff tomorrow. The firm wishes to maximise the expected change in W_t^i plus their current payoff (Sannikov, 2008). Under the terms of the contract if the firm chooses abatement their continuation value will evolve according to $dW_t^i = \gamma^i W_t^i dt - (\pi^i - c^i)dt - dI_t^i + \beta_t^i \sigma dZ_t$. If they instead choose not to abate they gain the avoided cost of abatement c^i but their continuation value will evolve differently. The firm's failure to abate affects the drift of pollution but not the volatility. Hence their promised value will evolve as $dW_t^i = \gamma^i W_t^i dt - (\pi^i - c^i)dt - dI_t^i + \beta_t^i(A^i dt + \sigma dZ_t)$.[4] The regulator must ensure that the change to the firm's payoff from deviating (not abating at time t) is negative, that is:

$$\beta_t^i A^i dt + c^i dt \leq 0$$

The first term ($\beta_t^i A^i dt$) is the present value change in firm i's promised value from not abating. The second term is the immediate gain to the firm from not abating: the total cost of abatement

[4] Note that the regulator cannot observe the deviation and hence adjusts the promised value for the cost of abatement even if the firm shirks.

$c^i dt$. Re-arranging:

$$\beta_t^i \leq -\lambda^i$$

where $\lambda^i = \frac{c^i}{A^i}$, the unit cost of abatement. Hence for a positive pollution flow dY_t the firm's promised value must fall by at least λdY_t.

2.2 The incentive payments

Now that we know the form of dW_t^i we can proceed to characterise the HJB equation of the regulator's problem (the regulator's value function) and find the optimal contract. Define $\mathbf{W_t}$ as the set of continuation payoffs for all N_t firms at time t. We wish to characterise the function $b(\mathbf{W_t})$ which is the highest expected payoff to the regulator from the promised payoffs $\mathbf{W_t}$ and the stock of pollution Y_t. This will allow us to specify the optimal contract and incentive payments. Proposition 1 outlines their form.

Proposition 1. *If the regulator always implements the abatement action then the Hamilton-Jacobi-Bellman equation $b(\mathbf{W})$ is separable in W^1, W^2 ... W^N and is given by:*

$$b(\mathbf{W}) = \sum_i f^i(W^i) \tag{9}$$

where

$$\delta f^i(W^i) = \begin{cases} \tau^i + f^i_{W^i}(W^i)\left(\gamma^i W^i - (\pi^i - \lambda^i A^i) + \tau^i\right) + \frac{1}{2}\left(\sigma\lambda^i\right)^2 f^i_{W^i W^i}(W^i) \\ \qquad \text{for } R^i \leq W^i \leq \hat{W}^i \\ f^i(\hat{W}^i) - (W^i - \hat{W}^i) \\ \qquad \text{for } W^i > \hat{W}^i \end{cases} \tag{10}$$

and the incentive payment to the firm takes the form:

$$dI_t^i = \begin{cases} -\tau^i dt & W_t^i < \hat{W}^i \\ (W_t^i - \hat{W}^i)dt & W_t^i \geq \hat{W}^i \end{cases} \tag{11}$$

where the threshold $\hat{W}^i$ is the lowest value of W^i at which $f^{i'}(W^i) = -1$.

Proof. For the purposes of the proof we assert concavity of the function $f(.)$.[5] We first argue that the HJB is separable in W^i and W^j. Note that under the optimal action the evolution of the continuation value dW_t^i is correlated with dW_t^j through the pollution shock σdZ_t but W_t^i does not depend directly on W_t^j. The continuation values would be linked for example if the incentives or recommended action for firm i depended on the actions of firm j. However the recommended action for firm i is independent of the actions of firm j and independent of whether or not firm j is active.[6] The regulator therefore cannot gain by linking the payoffs of the firms and will choose to adjust dI^i and dW^i independently of dI^j and dW^j. But then the regulator's value function is separable in W_t^i and W_t^j, and we can write:

$$
\begin{aligned}
b(\mathbf{W_t}) &= -e^{\delta t} \sum_i \int_t^{T^i} e^{-\delta s} dI_s^i \\
&= \sum_i f^i(W^i)
\end{aligned}
$$

With separability we can focus on $\delta f^i(W^i)$ Using Ito's Lemma (see the appendix for details) we can write:

$$
\delta f^i(W_t^i)dt = -dI_t^i + f_{W^i}^i(W^i)(\gamma^i W_t^i dt - (\pi^i - \lambda^i A^i)dt - dI_t^i) + \frac{1}{2}(\sigma\beta^i)^2 f_{W^iW^i}^i(W_t^i)dt \qquad (12)
$$

Now note that the regulator can always choose to increase the firm's payoff by amount dI_t^i today and reduce the firm's continuation value by an amount that keeps the firm exactly indifferent (DeMarzo and Sannikov, 2006). Hence for the function f^i to be part of the value function it must be that $f^i(W_t^i - dI_t^i)dt - dI_t^i \leq f^i(W_t^i)dt$. That is, the actual value to the regulator must be greater than the value of shifting payments from now to the future. This implies that:

$$
f_{W^i}^i(W_t^i) \geq -1 \qquad (13)
$$

[5] We prove concavity in appendix C

[6] In the definition of b this means that the expected stopping time $E(T^j)$ and the incentives dI_t^j has no impact on the recommended action or incentive compatibility constraint of firm i for any t.

The regulator can also substitute an amount $dI_t^i < 0$ for an amount $dW_t^i > 0$. If the regulator was always able to make this substitution then $f^i(W_t^i + dI_t^i)dt + dI_t^i \geq f^i(W_t^i)dt$ and $f^{i'}(W^i) = -1$ everywhere. If $f_{W^i}^i(W_t^i) > -1$ anywhere then the regulator is better off increasing W^i and taking an amount dI^i now. But the regulator must set $dI^i \geq -\tau^i$. Then if $dI^i > -\tau^i$ it must be that $f_{W^i}^i(W^i) = -1$. Conversely, if $dI^i = -\tau^i$ then $f_{W^i}^i(W_t^i) > -1$.

Let $\hat{W}^i$ be the lowest value of W^i such that $f_{W^i}^i(\hat{W}^i) = -1$. Then, with f^i concave, this value defines a threshold in W^i. For $W_t^i < \hat{W}^i$, $f_{W^i}^i(W_t^i)dt > -1$ and $dI^i = -\tau^i$. Above $\hat{W}^i$ the regulator is better off rewarding the firm with a direct payment. Then it is optimal for the regulator to choose dynamic incentives for the firm dI^i such that:

$$dI_t^i = \begin{cases} -\tau^i dt & W_t^i < \hat{W}^i \\ (W_t^i - \hat{W}^i)dt & W_t^i \geq \hat{W}^i \end{cases}$$

Finally it remains to show that $\beta^i = -\lambda^i$. With f^i strictly concave in W^i, $f_{W^iW^i}^i(W^i) < 0$ and the regulator will seek to minimise the magnitude of β^i subject to the incentive compatibility constraint $\beta_t^i \leq -\lambda^i$. Hence $\beta^i = -\lambda^i$ and (12) becomes

$$\delta f^i(W^i) = \tau^i + f_{W^i}^i(W^i)(\gamma^i W^i - (\pi^i - c^i) + \tau^i) + \frac{1}{2}(\lambda^i \sigma)^2 f_{W^iW^i}^i(W^i) \tag{14}$$

□

Separability of the HJB **W** relies on the assumption that the recommended actions of active firms are independent of the total number of active or exited firms. Without this assumption, the termination time T^i may affect firm i''s recommended action and separability may fail. The proof also relies on an assumption of concavity which we prove in appendix C.

The HJB is thus a second order ordinary differential equation. To pin down the solution to $f^i(W^i)$ and the boundary $\hat{W}^i$ we require the following three conditions: $f^i(R^i) = 0$ (initial value condition), $f_{W^i}^i(\hat{W}) = -1$(the smooth pasting condition) and $f_{W^iW^i}^i(\hat{W}) = 0$ (the supercontact condition). The initial value condition says that at the outside option of firm i, the regulator's value $f^i(.)$ is zero. The smooth pasting and supercontact conditions impose continuity at the threshold value $\hat{W}^i$.

Lemma 1 and Proposition 1 outline the structure of the optimal regulatory contract and the dynamic incentives that induce each firm to abate. The regulator observes $\mathbf{W}$ and Y, and implements transfers according to the N_t contracts. These contracts charge firm i a constant amount τ^i and provide a performance payment $dI^i > 0$ once i's promised value exceeds a threshold $\hat{W}^i$. We can thus interpret τ^i as a heterogeneous lump sum tax or a license fee. Once the continuation value of the firm exceeds $\hat{W}^i$ then the regulator pays them an amount dI^i to exactly bring them back to the threshold value (i.e. $W^i - \hat{W}^i$). That is, the size of the direct payment brings firm i back to the point where the regulator is once again indifferent between providing short and long run incentives. Hence the slope of the HJB in the region $W^i > \hat{W}^i$ is exactly -1. For $W^i < \hat{W}^i$ the regulator motivates the firm by allowing W_t^i to increase. At the threshold $\hat{W}^i$ this changes, the regulator is exactly indifferent between providing incentives via a direct payment, or via an increase in the firm's promised value. Below the threshold if firm i abates, their promised value has drift $(\gamma^i W_t^i - (\pi^i - c^i) + \tau^i)dt$ and volatility $-\lambda^i \sigma dZ_t$.

The regulator's value from the contract $f^i(W^i)$ is concave in the continuation value of the firm. As W^i approaches R^i the risk of the firm exiting is higher and this reduces the regulator's expected value. Once the firm exits, the regulator no longer collects τ^i. As W^i approaches $\hat{W}^i$ the likelihood that the firm will receive an incentive payment increases, this also reduces the regulator's expected future tax receipts. To minimise the probability of exit, the regulator 'back loads' the firm's performance payments. Below the threshold $\hat{W}^i$ the optimal contract allows firm i's promised value W_t^i to grow as a reward for abatement. As W_t^i grows the probability that i exits decreases. This is valuable to the regulator as it only collects net transfers when the firm operates. We now go on to explore how this stock, or the regulator's value, varies with the parameters of the model.

2.3 Comparative statics

We wish to understand how the regulator's value $f^i(W^i)$ changes with the parameters of the model. Unlike $h(Y)$ we have no closed form solution for $f^i(W^i)$. The value $f^i(W^i)$ represents the expected future accumulation of tax receipts from the firm. Following DeMarzo and Sannikov (2006) (see their Lemmas 4 and 6), using the envelope theorem it is simple to derive an expression for the change in the regulator's value as a result of a change in one of the parameters. This allows us to derive Lemma 2.

Lemma 2. *The regulator's value is weakly increasing in the maximum transfer τ^i, is decreasing in the volatility of pollution σ. It is also decreasing in the outside option of the firm and weakly* *decreasing in the discount rate of the firm.*

Lemma 2 is proven in appendix D. The regulator's value $f^i(W^i)$ is thus increasing in the payments τ^i and decreasing in the volatility of pollution (consistent with the concavity of the regulator's value), the firm's outside option, and in the discount rate of the firm. The intuition for these results is fairly clear. Holding all else constant, an increase in the parameter τ^i implies greater transfers to the regulator. The volatility of pollution makes the moral hazard problem harder to overcome. A higher outside option implies that the firm is more likely to exit inefficiently. Finally, the more impatient the firm the less ability the regulator has to defer payments. This means that incentives must be brought forward into the present which is costly to the regulator. We are not able to sign the comparative statics with respect to the firm's profits, the unit cost of abatement λ^i or the abatement level A^i.

2.4 Optimality of abatement

Thus far we have characterised the contract under the assumption that all active firms abate. When a firm does not abate ($a^i = 0$) the stock of pollution evolves according to $dY = A^i + \sigma dZ$ where A^i is the expected emission level of firm i given no other firm pollutes. If the regulator allowed the firm to cease abatement for a period then in that period it no longer needs to provide the firm with incentives or a subsidy to cover the cost of abatement. Hence if the regulator allowed the firm to emit then the evolution of the firm's continuation value could be independent of the stock of pollution (i.e. instead of providing incentives to the firm by setting $\beta^i = \lambda^i$ the regulator sets $\beta^i = 0$). This would also reduce the volatility of the regulator's payoff, which is valuable to them (their payoff function is concave in W^i as shown in appendix C). Hence two processes for the firm's continuation value are possible, depending on whether the the firm is allowed to cease abatement. These are:

$$dW^i = \gamma^i W^i - (\pi^i - \lambda^i A^i) - dI^i + \lambda^i \sigma dZ, \;\; \beta^i = \lambda^i \tag{15}$$

$$dW^i = \gamma^i W^i - (\pi^i - \lambda^i A^i) - dI^i, \;\; \beta^i = 0 \tag{16}$$

Note that W^i follows (16) only when the regulator *allows* the firm to cease abatement. If the regulator expects the firm to abate but they fail to do so then the firm's continuation value will move in line with the stock of pollution. The following proposition summarises the conditions for optimality of abatement when the (flow) damage from pollution at time t is equal to Y_t, the stock of pollution at time t:

Lemma 3. *If:*

$$\frac{p(\mathbf{N_t})}{\delta} \geq -\sum_i b_{W^i}(\mathbf{W})\lambda^i A^i - \frac{1}{2}(\sigma\lambda^i)^2 b_{W^iW^i}(\mathbf{W}) \tag{17}$$

then abatement by firm i is optimal.

Proof is in the appendix. To see the intuition behind this constraint first note that the left-hand side of 17 represents the social cost of pollution, the right hand side is the cost of abatement and the cost of providing incentives such that firms abate. Lemma 3 thus suggests that abatement is optimal $\forall(\mathbf{W}, Y)$ if the costs of pollution are greater than the costs of abatement plus the costs of providing the firm with incentives. Second, note that $b_{W^i}(\mathbf{W}) \geq -1$ hence

$$-\sum_i b_{W^i}(\mathbf{W})\lambda^i A^i - \frac{1}{2}(\sigma\lambda^i)^2 b_{W^iW^i}(\mathbf{W}) \leq \lambda^i A^i - \frac{1}{2}(\sigma\lambda^i)^2 b_{W^iW^i}(\mathbf{W})$$

The term $\frac{1}{2}(\lambda^i\sigma)^2 b_{W^iW^i}(\mathbf{W})$ thus represents the costs to the regulator of providing both incentives and insurance to the firms; it is increasing in the cost of abatement and the volatility of pollution.

In a first best world pollution is optimal if $\frac{p(\mathbf{N})}{\delta} \geq \lambda^i A^i$. As $\delta < 1$ abatement may be optimal even if $A^i < \lambda^i A^i$. This is because the regulator suffers damage from the stock of pollution not the flow. As $b_{WW}(\mathbf{W}) < 0$ abatement is optimal less often in the presence of moral hazard. Moreover, this constraint may bind as the volatility of pollution (σ) increases or as the required incentives (λ^i, which is also the unit cost of abatement) increase.

3 Budget balance and the tax rate τ

3.1 The performance threshold

The performance threshold represents the point at which the regulator switches between providing the firm with short run and long run incentives. In other words, in the region below $\hat{W}^i$ the regulator motivates the firm by making promises about future payoffs; the firm continues to pay tax τ^i which is collected by the regulator. Once the firm's value exceeds $\hat{W}^i$ the regulator returns a portion of these tax receipts as an incentive payment. The threshold $\hat{W}^i$ represents the point at which the regulator is indifferent between providing short run and long run incentives. In essence, for $W^i \leq \hat{W}^i$ it is cheaper to reward firm i via the promise of future payments than to give a one off tax cut or subsidy in the present. Above $\hat{W}^i$ it is cheaper to reward the firm via a payment.

We now proceed to derive the value of $\hat{W}^i$. We assume that the regulator must on average balance their budget. Define expected budget balance as $\sum_i f^i(W_0^i) = 0$ where W_0^i is the initial value of firm i. If this value is below the firm's reservation value then the firm immediately quits. With an initial value above the firm's reservation level the firm accepts the contract and produces. We assume a firm's starting value is equal to their average expected profits minus the cost of abatement, which we refer to as the social value of the firm's production. That is, $\gamma^i W_0^i = \pi^i - \lambda^i A^i$. This allows us to specify the threshold $\hat{W}^i$ which we do in the following proposition:

Proposition 2. *If the regulator balances its budget and the firms have initial value* $\mathbf{W_0}$ *equal to their average social value of their production, then each firm's reflecting boundary* $\hat{W}^i$ *is also defined by the social value of the firm's production i.e.* $W_0^i = \hat{W}^i = \frac{\pi^i - \lambda^i A^i}{\gamma^i}$. *Moreover budget balance determines the transfer* τ^i.

Proof. It can be shown that

$$\sum_i \delta f^i(W^i) + \gamma^i W^i - (\pi^i - \lambda^i A^i) \leq 0 \tag{18}$$

with equality at the threshold value $\hat{W}^i$. Substituting the budget balance condition ($\sum_i f^i(W_0^i) = 0$) and the firm's starting values ($\gamma^i W_0^i = \pi^i - \lambda^i A^i$) into the left hand side of (18):

$$\delta \sum_i f^i(W_0^i) + \gamma^i W_0^i - (\pi^i - \lambda^i A^i) = 0$$

Hence it must be that $\hat{W}^i = W_0^i$. To see this note that the left hand side of (18) is increasing in W^i because $f^{i'} \geq -1$ and $\gamma^i > \delta^i$. Hence $\hat{W}^i = W_0^i$ is only value of W^i that satisfies (18) with equality.

If the regulator does not budget balance its budget there are unknowns ($f^i_{W^i}(.), f^i_{W^iW^i}(.)$ and $\hat{W}^i$) which are solved from the conditions $f^i(R) = L$, $f^i_{W^i}(\hat{W}^i) = -1$ and $f^i_{W^iW^i}(\hat{W}^i) = 0$. With budget balance if τ^i is given, then we have an over-identified system. With budget balance we have the three unknowns: $f^i_{W^i}(.), f^i_{W^iW^i}(.)$ and τ^i for which we have three conditions: $f^i(R) = L$, $f^i_{W^i}(\hat{W}^i) = -1$ and $f^i_{W^iW^i}(\hat{W}^i) = 0$ where $\hat{W}^i$ for $i = 1..N$ are known.

□

The threshold represents the point at which the regulator decides to motivate firms with a short run payment. Below this threshold the firms pay a tax, above the threshold they receive a tax rebate. Firms' actions, along with the volatility of pollution, affect their rating W^i and hence their promised value. When a firm's rating exceeds its social value the firm is paid out. Under the abatement action, a firm's promised value will move randomly around its social value and the regulator, in expectation, does not collect any net tax. A firm enters with promised value equal to their social value and in expectation each entering firm will receive the social value of their production. If a firm does not abate, in expectation their promised value will fall below their social value, they will pay higher net taxes and the regulator's value will have positive expectation.[7]

As we have an analytical solution for each boundary $\hat{W}^i$ we are able to explicitly study comparative statics with respect to a firm's threshold. Higher profits will increase the boundary value by a factor of $\frac{1}{\gamma^i}$ whilst increases in either the cost of abatement λ^i or the level of abatement (A^i) reduces firm i's threshold by the same magnitude. The volatility of pollution and the discount rate of the regulator have no effect on the threshold. The volatility of pollution affects the firm's payoffs via the process dW^i, the threshold determines the optimal mix of short run versus long run incentives. The threshold decreases with a more impatient firm because it is more costly to postpone payment to an impatient firm. The optimal balance of long run and short run incentives must therefore change.

3.2 The firms' outside option

In the optimal contract the dynamics of the incentives are important. The regulator uses future promises to the firms to reward abatement and punish pollution. This means that the firms' continuation values evolve with shocks to pollution, and introduces the risk that a firm exits if,

[7] In the model without limited liability budget balance is similarly achieved if $W^i_0 = \frac{\pi^i - \lambda^i A^i}{\gamma^i}$.

after a series of positive shocks to pollution, their future promised value does not exceed their outside option. The contract is designed to minimise this possibility. The firm's exit is inefficient as the firm's profit net of the cost of abatement is positive. But the threat of exit is important to a firm's incentives to abate in the vicinity of the outside option. If the firms shirk the lowest instantaneous payoff that they can receive is $\pi^i - \tau^i$. If they receive this payoff forever then their continuation value $W^i = \frac{\pi^i - \tau^i}{\gamma^i}$. To ensure that the pollution stock remains constant, the regulator must be able to force a firm to exit or abate. That is, every firm must always prefer to exit or abate. But the regulator cannot push the promised payoff W^i below $\frac{\pi^i - \tau^i}{\gamma^i}$. To preserve the firms incentives this means that $R^i \geq \frac{\pi^i - \tau^i}{\gamma^i}$. This in turn implies lower bounds on τ, which are summarised in the following lemma.

Lemma 4. *The tax rate τ^i must satisfy two constraints: $\tau^i > \pi^i - \gamma^i R^i$ and $\tau^i > \lambda^i A^i$.*

Proof. We first show that if $R^i \leq \frac{\pi^i - \tau^i}{\gamma^i}$ then the firm will not exit. This occurs if $W_t^i \geq R^i$. Hence if the firm does not exit but abates then:

$$E_t \left[\exp^{\gamma^i t} \int_t^{T^i} \exp^{-\gamma^i s} d\Pi_s^i + \exp^{-\gamma^i (T^i - t)} R^i \right] \geq R^i$$

$$E_t \left[\exp^{\gamma^i t} \int_t^{T^i} \exp^{-\gamma^i s} \pi^i ds - dI_s + \exp^{-\gamma^i (T^i - t)} R^i \right] \geq R^i$$

$$E_t \left[(1 - \exp^{-\gamma^i (T^i - t)}) \frac{\pi^i}{\gamma^i} + \exp^{\gamma^i t} \int_t^{T^i} \exp^{-\gamma^i s} dI_s^i \right] \geq E_t \left[(1 - \exp^{-\gamma^i (T^i - t)}) \right] R^i$$

$$\frac{\pi^i}{\gamma^i} + E_t \left[\frac{\exp^{\gamma^i t}}{(1 - \exp^{-\gamma (T^i - t)})} \int_t^{T^i} \exp^{-\gamma^i s} dI_s^i \right] \geq R^i$$

$$\text{Now if:} \quad R^i > \frac{\pi^i - \tau^i}{\gamma^i}$$

$$\text{then:} \quad R^i > \frac{\pi^i}{\gamma^i} + E_t \left[\frac{\exp^{\gamma^i t}}{(1 - \exp^{-\gamma^i (T^i - t)})} \int_t^{T^i} \exp^{-\gamma^i s} dI_s^i \right]$$

for some realised paths of dI_s^i.

Now $R^i > \frac{\pi^i - \tau^i}{\gamma^i}$ implies that $\tau^i > \pi^i - \gamma^i R^i$. For the remainder of the lemma, note that to satisfy the firm's IR constraint, it must be that $W_0^i \geq R^i$, under budget balance $W_0^i = \frac{\pi^i - \lambda^i A^i}{\gamma^i}$, which gives the result that $\tau^i > \lambda^i A^i$. □

Lemma 4 places constraints on the value of the parameter τ^i. Specifically, it must be greater than the total cost of firm i's abatement, and it must be greater than the difference between the firm's profit flow π^i and the average value of the outside option $\gamma^i R^i$. We interpret $\pi^i - \gamma^i R^i$ as the marginal (private) benefit of operating. Absent the costs of abatement, it is the additional benefit flow that firm i receives from producing versus taking its outside option. To ensure that the firm will exit after some histories, and preserve its incentives around the outside option, the regulator must ensure that the flow from polluting forever is lower than the average value of the outside option. Without bearing the costs of abatement it must be the case that $\pi^i > \gamma^i R^i$ otherwise the firm would not be in business. The tax levied by the regulator must ensure that sometimes the firm prefers to exit over shirking, so the amount that the regulator takes away must be greater than marginal benefit the firm receives from producing.

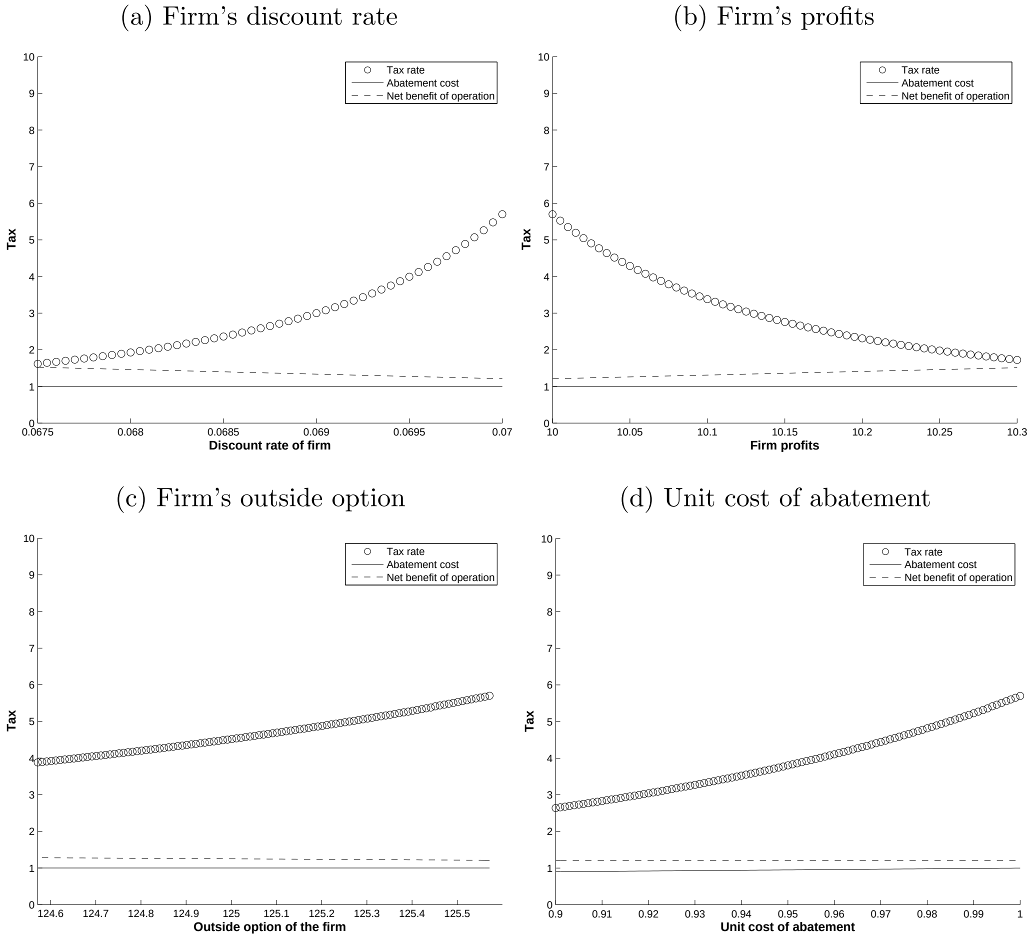
3.3 Numerical comparative statics for τ

As with the HJB, the tax rate has no closed-form solution. Our HJB now becomes a boundary value problem with unknown parameters τ^i. To explore how the tax rate varies with parameters of the model we numerically solve the regulator's HJB equation and the unknown parameters τ^i. Figures 1 and 2 show computed tax rates for specified parameter values. Each panel displays how the tax parameter τ^i as we vary a single parameter of the model and hold all other parameters constant. Panels (a) and (b) show how the tax rate varies with the firm's discount rate γ and the firm's profits π. Panels (c) and (d) show how the tax rate varies with the firm's outside option and the cost of abatement (R and λ respectively).

Panel (a) of 1 shows that the tax rate increases as the firm becomes more impatient (i.e. $\gamma \to 1$) and decreases as the firm earns greater profit. Note that γ^i changes the drift of W^i as well as the threshold $\hat{W}^i$. A more impatient firm has a lower threshold $\hat{W}^i$ and its promised value W^i grows at a faster rate. All other things equal, the regulator must collect greater revenue from the firm in order to be able to compensate an impatient firm more often and still maintain budget balance. Panel (b) shows that the tax rate decreases with the firm's profits. The firm's profits slow the growth of W^i and increase the threshold $\hat{W}^i$, hence they have the opposite effect to increases in impatience.

Figure 1 also demonstrates that the tax rate is an increasing function of the firm's outside option (Panel (c)) and the unit cost of abatement (Panel (d)). Increases in the firm's outside option make exiting more attractive to the firm. By taxing the firm more in the present, and increasing what is owed to the firm in the future, the regulator can reduce the probability that the

Figure 1: Computed tax rates: firm characteristics



Notes: Figures show computed tax rates for a range of parameter values. Baseline values for all calculations are $\gamma = 0.07, \pi = 10, \lambda = 1, A = 1, \sigma^2 = 4, \delta = 0.04, R = 125.5714$.

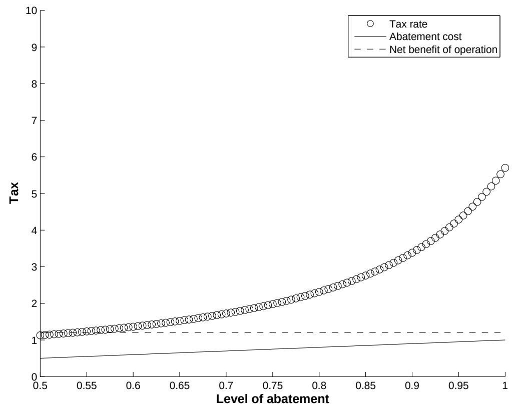
firm exits. Thus the regulator chooses to increase future promises i.e. increase W^i away from R^i by taking more away from the firm today.

Like γ and π the unit cost of abatement affects the threshold as well as the drift. Intuitively, increases in the cost of abatement lower the performance threshold. As the unit cost of abatement increases, the incentive of the firm to shirk is also greater. The regulator must provide a stronger incentive to abate, which is represented by a greater volatility of the firm's promised value. Greater volatility also means a higher change of larger payouts as the firm's promised value is more likely to exceed the performance threshold by a greater amount.

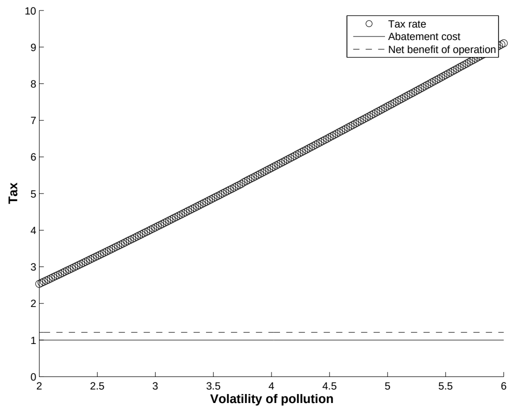
Figure 2 shows how the tax rate varies with the level of abatement A (panel (a)), the volatility of pollution σ (panel (b)) and the regulator's discount rate δ (panel (c)). The tax rate is increasing in the volatility of pollution and the level of abatement. The volatility of pollution affects the drift of the firm's promised value and not the performance threshold. An increase in volatility increases

Figure 2: Computed tax rates: pollution and regulator characteristics

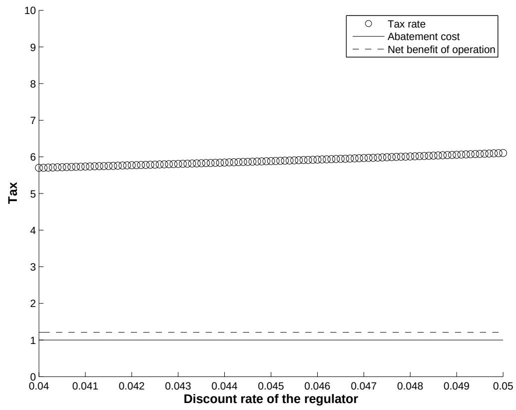
(a) Level of abatement



(b) Volatility of pollution



(c) Regulator's discount rate



Notes: Figures show computed tax rates for a range of parameter values. Baseline values for all calculations are $\gamma = 0.07, \pi = 10, \lambda = 1, A = 1, \sigma^2 = 4, \delta = 0.04, R = 125.5714$.

the probability of large payouts from a positive shock to pollution. Thus all else equal the regulator most collect a greater tax to ensure it meets budget balance. The tax rate also increases with the impatience of the regulator. The patience of the regulator does not affect the promised value of the firm or its performance threshold. A more impatient regulator has higher costs of providing rewards in the present. All else equal, a more impatient regulator would prefer to delay paying the firm. However it cannot as the evolution of W and the threshold value are fixed with respect to δ. Hence the tax rate τ must increase to compensate the more impatient regulator.

4 Comparing the contract

In this section we consider the design of the contract when the firm does not have limited liability and the regulator faces no constraints on what it can take from the firm in any instant.[8] This provides a benchmark to which we can compare the features of our contract.

Proposition 3. *If there are no constraints on the instantaneous transfer between the regulator and a firm, the regulator's value function* $\delta b(\mathbf{W}) = 0$*. A firm's promised value is constant. The regulator can ensure a firm abates by implementing dynamic transfers* $dI^i = -\lambda^i A^i - \lambda^i \sigma dZ_t$*. The instantaneous payoff of the firm is* $d\Pi^i = \pi^i - \lambda^i A^i - \lambda^i \sigma dZ_t$

Proof. We begin by specifying the form of dW^i and the incentives dI^i. In this model, Lemma 1 still holds:

$$dW_t^i = (\gamma^i W_t^i - (\pi^i - \lambda^i A^i)dt - dI_t^i + \beta_t^i dY_t \tag{19}$$

$$\beta_t^i \leq -\lambda^i \tag{20}$$

Following the logic of the proof for Proposition 1 the principal's value function $f^i(.)$ must have slope $= -1$ for all W^i. If $f^i_{W^i}(.) > -1$ then the regulator is better off allowing W^i to grow by taking away an amount dI^i from the firm. Without a limited liability constraint the regulator is always able to implement this transfer. At the solution then the regulator must always be exactly indifferent between providing the firm with current or future payoffs. Hence $f^i_{W^i}(.) = -1$ and that $f^i_{W^i W^i}(.) = 0$. Imposing this in (12):

$$\delta f^i(W_t^i)dt = (\pi^i - \lambda^i A^i)dt - \gamma^i W_t^i dt \tag{21}$$

But then $f^i_{W^i}(W) = -\frac{\gamma^i}{\delta} < -1$ and the regulator is better off providing an immediate transfer. To ensure that $f^i_{W^i}(.) = -1$ it must be that $dW^i = 0$. That is, the firm's continuation value is

[8] If we remove the policy constraint on the regulator and maintain limited liability of the firm it can be shown that the regulator sets $dI_t = -(l^i + (\pi^i - c^i))dt$ for $W_t^i < \hat{W}^i$ where $l^i dt$ represents the firm's exogenously given credit limit.

constant. Imposing this in (19) and assuming that the starting value $W_0^i = \frac{\pi^i - \lambda^i A^i}{\gamma^i}$:

$$dI_t^i = \beta_t^i dY_t \tag{22}$$

Then the regulator is indifferent to β^i subject to the incentive compatibility condition $\beta^i \leq -\lambda^i$.

□

We have shown that in an environment without limited liability, the optimal contract is a tax on the flow of pollution. Without limited liability, the firm can costlessly absorb unbounded variation in its instantaneous payoff. Hence there is no need in this environment to utilise future promises to reward the firm. Incentives are provided purely via instantaneous taxes and subsidies of amount $\beta^i dY_t \leq -\lambda^i dY_t$. As long as the regulator offers the firm $W_0^i > R^i$ the firm accepts the contract. The incentives for the firm to abate are purely provided in the short run. As $dW^i = 0$ the firm's continuation value remains at $W_t^i = W_0^i$ for all t. Without limited liability the future expected payoff does not change and hence the firm does not exit. Thus with no cost to volatility, the dynamics of the contractual relationship are not as important as in the limited liability case.

Both the limited liability and the benchmark model tie the firm's payoff to the pollution shock σdZ_t. Tying the firm's payoff to pollution is necessary to provide the firm with incentives to abate. However doing so introduces volatility into the payoff of the firm. In the benchmark model this volatility is costless. In the limited liability model that is not the case. The contract with limited liability ties the firm's *continuation* value to the flow of pollution. This means that the firm is less exposed to the volatility of pollution in the present. The shock to pollution is absorbed into the firm's promised value; it only affects the firm's instantaneous payoff around the performance threshold and around the outside option. The contract that is constrained to satisfy limited liability motivates the firm with long run incentives, however to do so it must introduce the possibility that the firm will exit. There is no such risk in the benchmark no limited liability model.

The benchmark contract is similar to an ambient tax however with an important difference. In the standard ambient tax model the tax paid by a firm is a function of the current stock and not the shock to pollution as in our benchmark contract. In the benchmark model studied here the individual rationality constraint must be satisfied. However as long as $W_0^i > R^i$ there is no risk of the firm's exit because the firm's future payoff is fixed at $W_0^i > R^i$. Neither Xepapadeas (1992) nor Athanassoglou (2010), who consider ambient taxes in a dynamic setting, consider the possibility that the firm has an outside option and hence do not allow for the the possibility of firm exit. However in these models, following a large shock that drags the stock of pollution away from the target level, firms may expect to continue to pay large amounts for some time even whilst they abate. With a large-enough outside option a firm may thus be better off exiting following a large shock. Tying the firm's payoff to the pollution shock and not the current stock of pollution avoids this issue. Hence neither our benchmark nor the limited liability setting result in a contract that takes the form of an ambient tax.

5 Conclusion

Non point source pollution is difficult to model and to regulate, yet it causes significant environmental damage. This type of pollution is characterised by hidden actions and random observed levels of pollution. Proposed solutions to the regulatory problem include input taxes and taxes on the stock of pollution. Neither have been widely implemented in practice. In contrast to the literature, we specify a rich model of firm incentives in a dynamic setting. We solve for the optimal contract with a regulator and multiple firms who suffer from limited liability and who each have outside options. The flow of pollution follows a brownian motion whose drift is affected by the actions of all firms. We find the contract for the case in which it is optimal for the regulator to keep pollution stock constant, and where the regulator must achieve expected budget balance. In this contract, each firm is initially promised the expected social value of their production. The regulator motivates the firm to abate by tying the firm's promised value to the realisation of pollution. Increases in the pollution stock cause the promised value to fall, whilst decreases in the stock cause the promised value to rise. If the firm abates, their promised value grows over time at their discount rate. Whilst the firm's promised value is below their social value they pay a constant tax. Once their promised value exceeds their social value, they are rewarded with a bonus payment. On average, if the firm abates, the regulator will achieve budget balance. We explore how the constant tax varies with the parameters of the model, and demonstrate how the contract differs when we relax the limited liability constraint.

References

ALBUQUERQUE, R. AND H. HOPENHAYN (2004): "Optimal lending contracts and firm dynamics," *The Review of Economic Studies*, 71, 285–315.

ATHANASSOGLOU, S. (2010): "Dynamic nonpoint-source pollution control policy: ambient transfers and uncertainty," *Journal of Economic Dynamics and Control*, 34, 2494–2509.

BIAIS, B., T. MARIOTTI, J. ROCHET, AND S. VILLENEUVE (2010): "Large risks, limited liability, and dynamic moral hazard," *Econometrica*, 78, 73–118.

CHAMBERS, R. AND J. QUIGGIN (1996): "Non point source pollution as a multi task principal agent problem," *Journal of Public Economics*, 59, 95–116.

CLEMENTI, G. AND H. HOPENHAYN (2006): "A theory of financing constraints and firm dynamics," *The Quarterly Journal of Economics*, 121, 229–265.

DEMARZO, P. AND Y. SANNIKOV (2006): "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *The Journal of Finance*, 61, 2681–2724.

EPA, U. (2009): "National Water Quality Inventory: Report to Congress," Tech. rep., Washington, DC: Environmental Protection Agency.

FISHER-VANDEN, K. AND S. OLMSTEAD (2013): "Moving Pollution Trading from Air to Water: Potential, Problems, and Prognosis," *The Journal of Economic Perspectives*, 27, 147–171.

KARP, L. (2005): "Nonpoint source pollution taxes and excessive tax burden," *Environmental and Resource Economics*, 31, 229–251.

LAFFONT, J. (1995): "Regulation, moral hazard and insurance of environmental risks," *Journal of Public Economics*, 58, 319–336.

SANNIKOV, Y. (2008): "A continuous-time version of the principal-agent problem," *The Review of Economic Studies*, 75, 957–984.

SEGERSON, K. (1988): "Uncertainty and incentives for nonpoint pollution control," *Journal of Environmental Economics and Management*, 15, 87 – 98.

Shortle, J. and R. Horan (2002): "The economics of nonpoint pollution control," *Journal of Economic Surveys*, 15, 255–289.

Spear, S. and S. Srivastava (1987): "On repeated moral hazard with discounting," *The Review of Economic Studies*, 54, 599–617.

WRI (2011): "World hypoxic and eutrophic coastal areas (map)," Tech. rep., World Resources Institute (WRI) accessed at http://www.wri.org/our-work/project/eutrophication-and-hypoxia on Nov 12, 2013.

Xepapadeas, A. (1992): "Environmental policy design and dynamic nonpoint-source pollution," *Journal of Environmental Economics and Management*, 23, 22 – 39.

——— (2011): *Annual Review of Resource Economics*, Annual Reviews, chap. The Economics of Non-Point-Source Pollution, 355–373.

Zhu, J. Y. (2013): "Optimal Contracts with Shirking," *The Review of Economic Studies*, 80, 812–839.

A Lemma 1: evolution of the firm's value

Proof. Define firm i's promised value $W_t^i(\mathbf{a})$ as the total expected payoff the firm receives from choosing abatement from time t after some history of reports and technology choice $(a_s^i, 0 \leq s \leq t)$, given that all other firms follow the recommended action from time t.

$$W_t^i(\mathbf{a}) = E_t \left[\int_t^{T^i} \exp^{-\gamma^i(s-t)} d\Pi_s^i(\mathbf{a}) + \exp^{-\gamma^i(T^i-t)} R^i \right]$$

Consider the case where for $0 \leq s \leq t$ the firm follows the recommended action and abates (i.e. $a_s^i = 1 \forall 0 \leq s \leq t$). It is sufficient to show that the proposition (i.e. (7)) holds for this case (DeMarzo and Sannikov, 2006). Define

$$V_t^i(\mathbf{a}) = \left[\int_0^t \exp^{-\gamma^i s} d\Pi_s^i(\mathbf{a}) + \exp^{-\gamma^i t} W_t(\mathbf{a}) \right]$$

where V_t^i is the total past and future expected payoff to the firm from action choice $a_s^i = 1 \ \forall \ \ s$. Then $V_t^i = E_t[V_{t+s}^i]$; the total past and future expected payoff from choosing the abatement action, where the expectation is taken over information available at time t, is constant over time. In other words V_t^i is a martingale. By the martingale representation theorem there is a process β_t^i such that $dV_t^i = \exp^{-\gamma^i t} \beta_t^i(\mathbf{a}) dY_t$. Differentiating V_t^i with respect to t, and using the martingale property:

$$\begin{aligned} dV_t^i &= \exp^{-\gamma^i t} d\Pi_t^i(\mathbf{a}) - \gamma^i \exp^{-\gamma^i t} W_t^i(\mathbf{a}) dt + \exp^{-\gamma^i t} dW_t^i(\mathbf{a}) \\ \exp^{-\gamma^i t} \beta_t^i(\mathbf{a}) dY_t &= \exp^{-\gamma^i t} d\Pi_t^i(\mathbf{a}) - \gamma^i \exp^{-\gamma^i t} W_t^i(\mathbf{a}) dt + \exp^{-\gamma^i t} dW_t^i(\mathbf{a}) \\ dW_t(\mathbf{a})^i &= \gamma^i W_t^i(\mathbf{a}) dt - d\Pi_t^i(\mathbf{a}) + \beta_t^i(\mathbf{a}) dY_t \end{aligned}$$

Now substituting for dY_t from (1) and $d\Pi_t^i$ from (3):

$$dW_t^i(\mathbf{a}) = \gamma^i W_t^i(\mathbf{a}) dt - (\pi^i - c^i) dt - dI_t^i + \beta_t^i \sigma dZ_t \tag{23}$$

□

B Using Ito's Lemma to derive $f^i(W^i)$

Following DeMarzo and Sannikov (2006) define:

$$G_t \equiv \int_0^t \exp^{-\delta s}(-\sum_i dI_s^i) + \exp^{-\delta t} b(\mathbf{W_t}) \tag{24}$$

where G_t is the total payoff to the regulator up to time t and from t. Differentiating G_t with respect to time and applying Ito's Lemma:

$$\begin{aligned} e^{\delta t} dG_t = &- \sum_i dI_t - \lambda A dt - \delta b(\mathbf{W_t}) dt + \\ &\sum_i b_W^i(\mathbf{W_t}) dW_t^i + \frac{1}{2} \sum_i (\sigma \beta^i)^2 b_{W^i W^i}(\mathbf{W_t}) dt + \sum_{i \neq j} b_{W^i W^j}(\mathbf{W_t}) dW^i dW^j \end{aligned}$$

The martingale approach then proceeds by recognising that under the optimal policy G_t is a martingale and hence dG_t has expectation zero. Substituting for (7) and assuming that the regulator implements abatement we can take expectations and solve for $\delta b(\mathbf{W_t}) dt$:

G_t is a martingale hence

$$\delta b(\mathbf{W_t})dt = -\sum_i dI_t^i + \sum_i b_W^i(\mathbf{W_t})dW_t^i +$$
$$\frac{1}{2}\sum_i(\sigma\beta^i)^2 b_{W^iW^i}(\mathbf{W_t})dt + \sum_{i\neq j} b_{W^iW^j}(\mathbf{W_t})dW^i dW^j$$

Imposing separability i.e. $b_{W^iW^j} = 0$ and evaluating each f under the optimal action, gives us the following:

$$\delta f^i(W_t^i)dt = -dI_t^i + f_{W^i}^i(W_t^i)(\gamma^i W_t^i dt - (\pi^i - c^i)dt - dI_t^i) + \frac{1}{2}(\sigma\beta^i)^2 f_{W^iW^i}^i(W_t^i) \tag{25}$$

C Concavity of the function $f^i(W^i)$

Proposition 4. *The function f^i is strictly concave for $W^i < \hat{W}^i$*

Proof. We first show that

$$\delta f^i(W^i) + \gamma^i W^i < \pi^i - c^i \tag{26}$$

holds. Condition (26) states that the average expected profits of the firm and the payoff to the regulator cannot exceed the social value of the firm's production. To verify this intuitive condition we use the fact that $f^{i'}(W^i) > -1, W^i < \hat{W}^i$ and that for $W^i > \hat{W}^i$, $f'^i(W^i) = -1$ and hence $f^{i''}(W^i) = 0$. Substituting this into the regulator's HJB equation, it must be that:

$$\delta f^i(\hat{W}^i) + \gamma^i \hat{W} = \pi^i - c^i \tag{27}$$

That is, the boundary of the firm's continuation value is the point at which the social value of production (profits minus abatement costs) is equal to the expected discounted value that the regulator and the firm derive from the regulated transfers.

We wish to evaluate $\delta f^i(W^i) + \gamma^i W^i$ for $W^i < \hat{W}^i$. From the definition of a derivative, for small Δ we have that $f(\hat{W}^i - \Delta) \approx f^i(\hat{W}^i) - \Delta f^{i'}(\hat{W}^i - \Delta)$. Now $f^{i'}(\hat{W}^i - \Delta) > -1$ hence

$f^i(\hat{W}^i - \Delta) < f^i(\hat{W}^i) + \Delta$. Then it must be the case that:

$$\begin{aligned}
\delta f^i(\hat{W}^i - \Delta) + \gamma^i(\hat{W}^i - \Delta) &< \delta(f^i(\hat{W}^i) + \Delta) + \gamma^i(\hat{W}^i - \Delta) \\
\delta f^i(\hat{W}^i - \Delta) + \gamma^i(\hat{W}^i - \Delta) &< \delta f^i(\hat{W}^i) + \gamma^i\hat{W}^i + \Delta(\delta - \gamma^i) \\
\delta f^i(\hat{W}^i - \Delta) + \gamma^i(\hat{W}^i - \Delta) &< \delta f^i(\hat{W}^i) + \gamma^i\hat{W}^i \\
\delta^i f^i(\hat{W}^i - \Delta) + \gamma^i(\hat{W}^i - \Delta) &< \pi^i - c^i
\end{aligned}$$

where the second last inequality comes from the fact that $\gamma^i > \delta$ and the last uses (27). Hence $\delta f^i(W^i) + \gamma^i W^i < \pi^i - c^i$ for $W^i < \hat{W}^i$.

This allows us to prove that $f^{i''}(W^i) < 0$ for $W^i < \hat{W}^i$. From (14):

$$f^{i''}(W^i) = \frac{2}{(\sigma\beta^i)^2}\left[\delta f^i(W^i) - \tau^i - f^{i'}(W^i)(\gamma^i W^i - (\pi^i - c^i) + \tau^i)\right] \tag{28}$$

by (13) :

$$f^{i''}(W^i) \leq \frac{2}{(\sigma\beta^i)^2}\left[\delta f^i(W^i) - \tau^i + \gamma^i W^i - (\pi^i - c^i) + \tau^i\right]$$

by (26) :

$$\delta f^i(W^i) - (\pi^i - c^i) + \gamma^i W^i < 0$$

hence $f^{i''}(W^i) < 0$. Hence for $W^i < \hat{W}^i$ the function f^i is concave. □

D Lemma 2 : comparative statics

Proof.

$$\frac{\partial f^i(W^i)}{\partial \tau^i} = E\left[\int_0^{T^i} e^{-\delta t}(1 + f^{i'}(W_t^i))dt\right] > 0 \tag{29}$$

$$\frac{\partial f^i(W^i)}{\partial \sigma} = E\left[\int_0^{T^i} e^{-\delta t}(\sigma \lambda^{i2} f''(W_t^i))dt\right] < 0 \tag{30}$$

$$\frac{\partial f^i(W^i)}{\partial R^i} = -f^{i'}(R^i)e^{-\delta T^i} < 0 \tag{31}$$

$$\frac{\partial f^i(W^i)}{\partial A^i} = E\left[\int_0^{T^i} e^{-\delta t} f^{i'}(W_t^i)\lambda^i dt\right] \tag{32}$$

$$\frac{\partial f^i(W^i)}{\partial \lambda^i} = E\left[\int_0^{T^i} e^{-\delta t} f^{i'}(W_t^i)A^i + \sigma^2 \lambda f^{i''}(W_t^i)dt\right] \tag{33}$$

$$\frac{\partial f^i(W^i)}{\partial \pi^i} = E\left[\int_0^{T^i} -e^{-\delta t} f^{i'}(W_t^i)dt\right] \tag{34}$$

The signs of (29)-(31) follow. For the sign of $\frac{\partial f^i(W^i)}{\partial \gamma^i}$ we repeat the argument of DeMarzo and Sannikov (2006). Suppose that the firm's true discount rate was $\gamma^{i'} < \gamma^i$ but they were offered the contract designed for a firm with discount rate γ^i. Then the firm is better off as their continuation value grows at a higher rate than under the contract with their true time preference whilst the regulator has the same value as under γ^i. However the contract is not optimal for the firm with discount rate γ^i, the regulator can reduce the value to the firm whilst preserving their participation and their incentives. Hence it must be that the value to the regulator increases as the discount rate of the firm decreases. The result follows. □

E Lemma 3: Optimality of abatement

Proof. For abatement to be optimal $\forall \mathbf{W}$ we need to ensure that the damage caused by pollution exceeds the costs of implementing abatement $\forall \mathbf{W}$. Let $B(\mathbf{W}, Y)$ be the value from delivering $\mathbf{W}$ to firms at pollution stock Y. We assume that the damage of pollution at time t is equal to the stock of pollution Y_t. Then the value of the contract under abatement is $b(\mathbf{W}) - \frac{Y}{\delta}$.

Let $\widehat{b}(\mathbf{W}, Y)$ be the value from allowing firms to pollute at $(\mathbf{W}, Y)$. Then $B(\mathbf{W}, Y) = \max\left\{b(\mathbf{W}) - \frac{Y}{\delta}, \widehat{b}(\mathbf{W}, Y)\right\}$. To construct $\widehat{b}(\mathbf{W}, Y)$ we redefine G_t in (24) assuming that the regulator wishes to implement pollution and accounting for the damage of pollution.

$$\hat{b}(\mathbf{W_t}, Y_t) = -E_t\left[e^{\delta t}\int_t^{\infty} e^{-\delta s}Y_s ds + e^{\delta t}\int_t^{T^i}\cdots\int_t^{T^N}\sum_i e^{-\delta s}dI_s^i\right]$$

If pollution is the optimal policy G_t is a martingale and hence we can solve for $\widehat{b}(\mathbf{W}, Y)dt$. Note that if all firms always pollute then the value function is once again separable in all state variables and firms never exit (i.e. $\mathbf{M_t} = \emptyset$). Using the martingale approach and imposing separability:

$$\begin{aligned}\delta\widehat{b}(\mathbf{W_t}, Y_t)dt &= -Y_t dt + \widehat{b}_Y(\mathbf{W_t}, Y_t)dY_t + \sigma^2\widehat{b}_{YY}(\mathbf{W_t}, Y_t)dt \\ &\quad -\sum_i dI_t + \sum_i \widehat{b}_W^i(\mathbf{W_t}, Y_t)dW_t^i + \frac{1}{2}\sum_i(\sigma\beta^i)^2\widehat{b}_{W^iW^i}(\mathbf{W_t}, Y_t)dt\end{aligned}$$

Then $B(\mathbf{W}, Y) = \max\left\{b(\mathbf{W}) - \frac{Y}{\delta}, \widehat{b}(\mathbf{W}, Y)\right\}$. So for abatement to be optimal $\forall(\mathbf{W}, Y)$ i.e. $B = b$ with $\beta^i = 0\ \forall\ i$

$$\begin{aligned}\delta B(\mathbf{W_t}, Y_t)dt \geq{}& -Y_t dt + B_Y(\mathbf{W_t}, Y_t)p(\mathbf{N_t})dt + \sigma^2 B_{YY}(\mathbf{W_t}, Y_t)dt \\ & -\sum_i dI_t + \sum_i B_W^i(\mathbf{W_t}, Y_t)\left(\gamma^i W^i - \pi^i - dI^i\right)dt\end{aligned} \tag{35}$$

$$\begin{aligned}\delta B(\mathbf{W_t}, Y_t)dt \geq{}& -Y_t dt - \frac{p(\mathbf{N_t})}{\delta}dt + \sigma^2 B_{YY}(\mathbf{W_t}, Y_t)dt \\ & -\sum_i dI_t + \sum_i B_W^i(\mathbf{W_t}, Y_t)\left(\gamma^i W^i - \pi^i - dI^i\right)dt\end{aligned} \tag{36}$$

as $B_Y = -\frac{1}{\delta}$, $B_{YY} = 0$. If this is the case then $B(\mathbf{W}, Y_t) = b(\mathbf{W}) - \frac{Y}{\delta}$, but then we can re-write the condition as:

$$\delta b(\mathbf{W_t})dt \geq -\frac{p(\mathbf{N_t})}{\delta}dt + \sum_i \tau^i dt + \sum_i b_{W^i}(\mathbf{W_t}, Y_t)(\gamma^i W_t^i - \pi^i + \tau^i)dt \tag{37}$$

as $dI^i = -\tau^i$. Both sides are now proportional to dt and hence we have that abatement is always optimal if and only if

$$\delta b(\mathbf{W}) \geq -\frac{p(\mathbf{N_t})}{\delta} + \sum_i \tau^i + \sum_i b_{W^i}(\mathbf{W})(\gamma^i W^i - \pi^i + \tau^i) \tag{38}$$

If this condition holds $\forall W$ then abatement is always optimal.[9] Substituting into (38) for $\delta b(\mathbf{W})$ and re-arranging gives:

$$\frac{p(\mathbf{N_t})}{\delta} \geq -\sum_i b_{W^i}(\mathbf{W})\lambda^i A^i - \frac{1}{2}(\sigma\lambda^i)^2 b_{W^i W^i}(\mathbf{W}) \tag{39}$$

□

[9]This is the equivalent of equation (23) in DeMarzo and Sannikov