

# UC Berkeley

## CUDARE Working Papers

### Title

Dynamic Quotas with Learning

### Permalink

<https://escholarship.org/uc/item/88x3f17p>

### Authors

Karp, Larry  
Costello, Christopher J

### Publication Date

2000-10-18

ADAMS  
914  
DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY  
DIVISION OF AGRICULTURE AND NATURAL RESOURCES  
UNIVERSITY OF CALIFORNIA AT BERKELEY

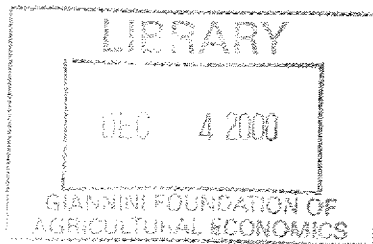
---

WORKING PAPER NO. 914

DYNAMIC QUOTAS WITH LEARNING

by

Larry Karp and Christopher Costello



---

Copyright © 2000 by Larry Karp and Christopher Costello. All rights reserved. Readers may make verbatim copies of this document for noncommercial purposes by any means, provided that this copyright notice appears on all such copies.

---

---

California Agricultural Experiment Station  
Giannini Foundation of Agricultural Economics  
October 18, 2000

# Dynamic Quotas with Learning

Christopher Costello

University of California, Santa Barbara

Larry Karp

University of California, Berkeley

October 18, 2000

## Abstract

We study the optimal quota sequence, in a stationary environment, where a regulator and a non-strategic firm have asymmetric information. The regulator is able to learn about the unknown cost parameter by using a quota that is slack with positive probability. It is never optimal for the regulator to learn gradually. In the first period, he either ignores the possibility of learning, or he tries to improve his information. Regardless of the outcome in the first period, he never experiments in subsequent periods.

*Keywords:* quotas, asymmetric information, searching

*JEL classification numbers* D83, L50

# 1 Introduction

An extensive literature compares taxes and quotas when firms and the regulator have asymmetric information about abatement costs. Following Weitzman (1974), many papers (including Malcomson (1978), Roberts and Spence (1976), Stavins (1996), Watson and Ridker (1984) and Yohe (1977)) compare the policies when damages are caused by the flow of pollution. More recently, a number of papers (Hoel and Karp (1998), Hoel and Karp (2000), Newell and Pizer (1998), Karp and Zhang (1999), Karp and Zhang (2000)) study the case where damages are caused by the stock rather than the flow of pollution. All of these models assume that the optimal quota is binding with probability 1. When quotas rights are traded efficiently (and firms are heterogenous, so that trade occurs), this assumption means that the equilibrium quota price conveys to the regulator the same information about an industry-wide cost shock as does the aggregate equilibrium response to the tax. If emissions trading does not occur – either because it is forbidden or too costly, or because homogenous firms have no incentive to trade – the assumption that the quota is always binding means that the regulator learns nothing about costs. In the latter case, quotas are obviously less informative than taxes.

The assumption that the optimal quota is binding with probability 1 is convenient because it focuses the model on central issues. In many cases, it is also reasonably descriptive of actual pollution problems. However, the assumption means that the models cannot be used to ask how a regulator might choose quotas to learn about abatement costs. This note shows how quotas can be used to acquire information about abatement costs in the absence of quota trading.

We use an infinite horizon model of a representative polluting firm. In each period the firm pollutes at its privately optimal level if this level is less than the quota; otherwise, the quota is binding. In each period, given his current beliefs about the cost function, the regulator chooses the quota in order to balance two conflicting objectives. He would like to control the amount of pollution, and simultaneously learn about the true costs so that he can choose a more efficient quota in the future. We assume that if the regulator uses the quota that would be optimal in a one period problem (the “myopic quota”), he learns nothing about costs. If he lets the firm produce at its privately optimal level, he learns everything about costs. We discuss these and other assumptions in the next section.

In each period, the regulator can use one of three types of quotas. He can use the myopic quota, in which case he maximizes welfare in the current period, but learns nothing about costs. He can use a (possibly infinite) quota which is slack with probability 1, in which case welfare

in the current period is low, but the regulator learns the true cost parameter. Finally, he can use a quota which is binding with probability strictly between 0 and 1. In this case, the regulator learns something about abatement costs, but expected welfare in the current period is lower than under the myopic quota.

One possibility is for the regulator to proceed cautiously. That is, he might refine his information over a number of periods until he eventually learns the value of the unknown parameter, or decides that further experimentation is too costly. Rob (1991) analyzes a model which has many of the characteristics of ours, and he finds that this kind of cautious approach is optimal. In Rob's setting, a social planner is uncertain about the location of the market demand curve.<sup>1</sup> The social planner gradually increases production capacity – requiring costly investment – until he learns the true market demand. In general, learning takes place for more than a single period.

Optimal learning in our setting is qualitatively different. We show that if there is any learning, it takes place in a single period. In the first period the regulator might decide to use any of the three types of quotas described above, in which case he either learns nothing with probability 1, he learns everything with probability 1, or he refines his information set but might or might not discover the true cost parameter. Whatever the outcome is in the first period, the regulator never experiments a second time.

The difference in optimal behavior in our search model and in Rob's is due to the difference in the costs of searching "cautiously" and "aggressively". A cautious search in Rob's model involves a slight increase in capacity. The amount of information obtained from such a search is modest, but so is the cost. Under the quota, however, it is necessary to incur a large cost in order to acquire even a small amount of information. This fact follows from the assumption that the myopic quota is binding with probability 1. An aggressive search in Rob's model has a higher expected cost than a more cautious search. For example, if the social planner builds twice the capacity needed to satisfy the market (in Rob's model), he incurs a greater loss than if he over-builds by a small amount. The cost of using a large quota, on the other hand, "levels off". The welfare cost of using a quota that is slack by a large margin equals the welfare cost of a quota that is slack by a small margin. These two differences make it unattractive (in the quota model) to use many cautious searches, but make it more attractive to use a single aggressive

---

<sup>1</sup>Much of Rob's analysis concerns the competitive equilibrium. However, for our purposes, the relevant material is his treatment of the social planner's problem.

search.

The next section presents the model and shows this result. We close with a brief discussion of the result.

## 2 The model and result

Here we list the assumptions of our model and explain their role. The next subsection uses an example to illustrate these assumptions and to describe the basic properties of the model. The following subsection shows, in the general setting, that learning occurs during at most one period.

We assume that the firm knows its abatement cost function, which is constant over time. There are neither exogenous nor endogenous (e.g. investment-related) cost changes. This assumption is not descriptive in many important pollution problems. However, it reduces the complexity of the problem, enabling us to understand how the regulator learns under quotas. We also assume that the firm does not behave strategically with respect to the regulator. Our model does not apply if strategic behavior is important. In many circumstances firms are genuinely non-strategic. In addition, we assume that social damages (external to the firm) are caused by the flow rather than the stock of pollution, and the damage function is constant. These assumptions eliminate two possible sources of dynamics, enabling us to focus on learning.

The regulator has full information about the cost function up to an unknown parameter  $\theta$ . This assumption reduces the dimension of the problem, and is standard in models of asymmetric information.

We define the myopic quota as the quota that would be optimal if the regulator ignores the possibility of learning; that is, the myopic quota minimizes the expectation of current dead-weight loss. In addition to the previous assumptions, we adopt:

**Assumption 1** *The firm's individually optimal level of emissions is monotonically increasing in its cost parameter,  $\theta$ .*

**Assumption 2** *The myopic quota is binding with probability 1.*

**Assumption 3** *(a) For quotas that are binding, the current loss in social welfare is a convex function of the quota. (b) If the quota is not binding, an increase in the quota does not affect*

*social welfare. (c) The welfare loss under symmetric information is 0.*

Assumption 1 states that the cost parameter affects the firm's individually optimal behavior. The regulator can calculate the smallest possible value of  $\theta$  consistent with a particular quota being exactly binding. The assumption also implies that if the quota is not binding, the regulator learns the true cost parameter. Thus, the regulator is certain to learn the cost parameter if he uses a quota that is slack with probability 1. (A later footnote gives an example of a situation where Assumption 1 is violated.)

If the firm is currently unregulated, then Assumption 1 means that the regulator knows the parameter  $\theta$ . In that case, there is obviously no asymmetric information. In most important cases, however, there exists some degree of regulation. The regulator may want to consider changing the current quota in order to improve information and to make it possible to use a more efficient quota in the future.

If the regulator is able to change the quota quickly, his problem is trivial. In that case, he does not regulate for a short period, learns the value of the unknown parameter, and then sets the efficient quota. Society may incur a large flow cost in the first period, but since it bears this cost for a short period it is unimportant. The search problem is interesting only if the regulator cannot change the quota arbitrarily quickly. We do not model the source of this sluggishness, but it appears to exist in many important situations.

Assumption 2 implies that under the myopic quota, the regulator never acquires any information. This assumption is consistent with the literature cited above, where there can be no learning under the optimal one-period quota.

Assumption 3 is standard. Part (a) implies that if the regulator ignores the possibility of learning, there is a unique optimal quota. If a quota is not binding, its value has no effect on the firm's action. An increase in the non-binding quota has no effect on deadweight loss, as part (b) states. Part (c) is merely a normalization. The deadweight loss of using a first-best (full information) policy is 0.

We explained in the Introduction that Assumptions 2 and 3b account for the qualitative difference in optimal searching behavior in this setting, compared to Rob's model. Assumption 2 means that the cost of acquiring a small amount of information is large, making a cautious search unattractive. Assumption 3b means that the cost of using a high quota levels off as soon as the quota is slack. The example below illustrates these features.

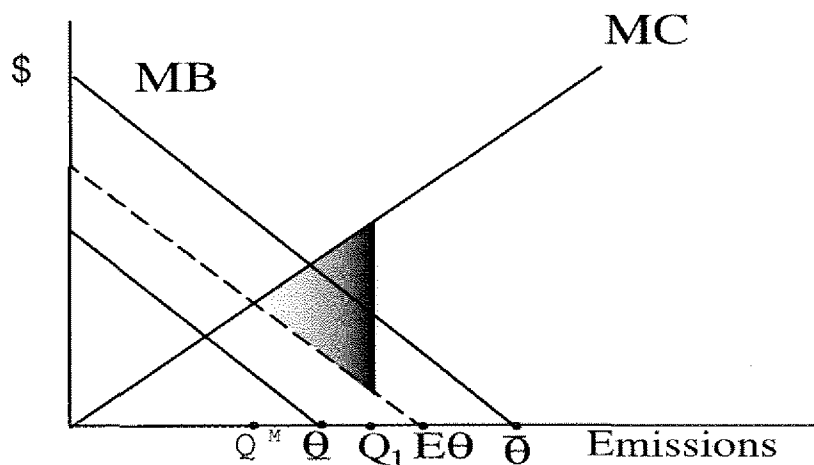


Figure 1: The linear example

## 2.1 The example

Here we illustrate the model for the case where the social marginal damage of pollution and the private marginal benefits of polluting (i.e., private marginal costs of abating) are linear. The regulator knows the slope but not the intercept of the marginal benefit curve. The intercept is distributed over a finite support. Much of the previous literature uses this linear-quadratic model with additive uncertainty.<sup>2</sup>

In Figure 1, the (known) marginal cost of pollution curve is  $MC$ . The marginal benefit of pollution (i.e., the negative of the firm's marginal abatement cost) is  $MB$ . The horizontal intercept of  $MB$  (denoted  $\theta$ ) is the level of pollution that the unregulated firm would choose. The regulator views  $\theta$  as an unknown parameter with support  $[\underline{\theta}, \bar{\theta}]$  and expected value  $E\theta$ . The dashed line shows the marginal benefit curve associated with  $E\theta$ . With this linear model, the myopic quota, denoted  $Q^M$ , occurs at the intersection of the marginal cost and the expected marginal benefit curve. Consistent with Assumption (2),  $Q^M < \underline{\theta}$ . Consistent with Assumption (1), the regulator learns the true value of  $\theta$  if he chooses a quota greater than or equal to  $\bar{\theta}$ . If the true value of  $\theta$  happens to be  $E\theta$ , and if for some reason the regulator uses a quota  $Q_1$ , the deadweight loss in the current period is shown by the shaded area.

<sup>2</sup>A modification of the example in this section illustrates a situation that violates Assumption 1. Suppose that the horizontal intercept of the firm's marginal benefit curve,  $\theta$ , is known, but the slope of the curve is private information. In this case, the use of a quota does not enable the regulator to learn about the cost parameter.



It is clear from this figure that the regulator would never choose a quota between  $Q^M$  and  $\underline{\theta}$ . Such a quota causes a current loss in expected welfare without offering any possibility of learning. If the regulator's subjective distribution does not have a mass point at  $\underline{\theta}$  he would not want to use a quota close to  $\underline{\theta}$ , since such a quota provides negligible information but causes a substantial current welfare loss. If he decides to learn about the cost parameter, he uses a quota that is slack with probability strictly greater than 0.

If, for example, the regulator uses  $Q_1 > \underline{\theta}$  (Figure 1), then two things can happen. If  $\theta < Q_1$  he learns the true value of  $\theta$  and is able to use the socially optimal quota in the next period; in this case, the future deadweight loss is 0. If  $\theta \geq Q_1$  he does not learn the true value of  $\theta$ , but he learns that the lower bound of the support is higher than he previously thought. He replaces the previous (subjective) lower bound of the support,  $\underline{\theta}$ , with  $Q_1$ . He updates his subjective probability of  $\theta$  using Bayes' Rule, improving his ability to use an efficient quota in the next period.

The next section shows that at most one search is optimal. It might be optimal never to search, and to always use the myopic quota. Alternatively, the regulator might want to search one time. In that case, it might be optimal to use a quota that is slack with probability 1. That strategy insures that the regulator learns the true value of  $\theta$ . The regulator might want to "search cautiously", i.e. to use a quota that is slack with probability strictly between 0 and 1. With the cautious search, the regulator improves his information, but might or might not learn the true value of  $\theta$ .

Figure 2 uses the previous linear example with uniform priors over  $\theta$  to show that any of these strategies might be optimal. The figure graphs the expected present discounted value of deadweight loss for different levels of the period  $t$  quota. We hold constant the discount rate (2%), the upper bound of the support ( $\bar{\theta} = 25$ ) and the slope of the firm's marginal benefit of emissions (1). The figure shows results for three sets of the remaining parameter values, the slope of the marginal damage function,  $a$ , and the lower bound of the support,  $\underline{\theta}$ . These three sets of value of  $(a, \underline{\theta})$  are (.05, 23) (the dotted curve), (1, 23) (the dashed curve) and (.4, 14) (the solid curve).

When marginal damages are low (the dotted curve), the regulator minimizes the expected loss by setting  $Q_t = \bar{\theta} = 25$ , learning the true value of  $\theta$  in one period. For a larger value of marginal damages, but the same amount of uncertainty (the dashed curve), the myopic quota of  $Q_t^* = 12.1$  minimizes losses. For an intermediate level of marginal damages and a much

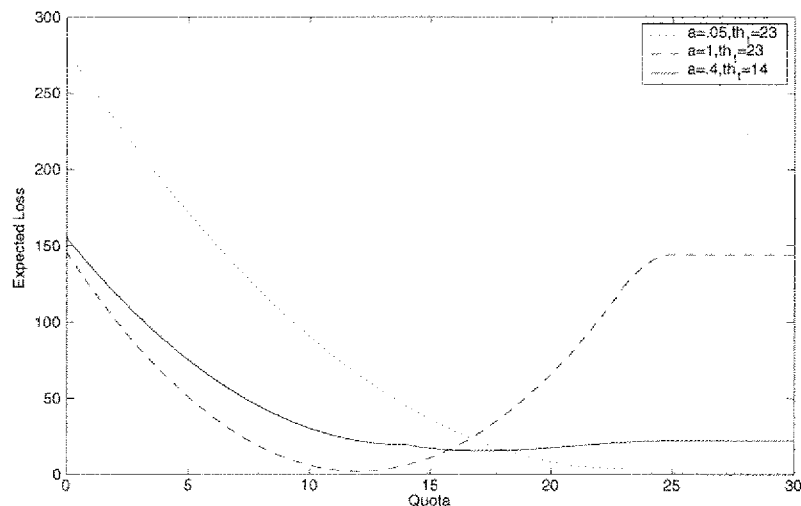


Figure 2: Expected loss as a function of the quota

higher level of uncertainty (the solid curve) the loss-minimizing quota equals 17.1, compared to the myopic quota of 13.9. The probability that the optimal quota is not binding is 0.28.

## 2.2 The general result

In this section we maintain the assumptions described above, but we allow the cost and benefit functions to be general. We do not need to specify the manner in which  $\theta$  affects private costs. However, in view of Assumption 1, we can treat the unknown parameter as the firm's privately optimal level of emissions (in the absence of regulation). This interpretation of the unknown parameter simplifies the proof below, and does not entail any loss in generality beyond Assumption 1. At time  $t$  the regulator treats  $\theta$  as an unknown parameter with support  $[\underline{\theta}_t, \bar{\theta}]$ . We assume that unregulated emissions are always finite, so  $\bar{\theta} < \infty$ .

If  $Q_t \leq \underline{\theta}_t$ , the regulator's subjective distribution over  $\theta$  in the subsequent period is unchanged. If  $\theta > Q_t > \underline{\theta}_t$  the regulator refines his information, calculating the subjective distribution in the subsequent period using Bayes Rule. In this case,  $\underline{\theta}_{t+1} = Q_t$ . In every period, the regulator can obtain the current subjective distribution using either the initial distribution or the previous distribution, together with the current value of the lower bound of the support. The lower bound summarizes all of the information that the regulator has acquired since the start of the program; it is the state variable in the regulator's optimal control problem.

We define  $L(Q, \theta)$  as the actual deadweight loss in the current period when the regulator uses the quota  $Q$  and the true value of the firm's unregulated level of emissions is  $\theta$ . We define  $Q^M(\underline{\theta}_t)$  as the myopic quota, i.e. the solution to

$$\min_Q E_t [L(Q, \theta)].$$

The operator  $E_t$  takes the expectation with respect to the unknown parameter  $\theta$ , given the regulator's beliefs at time  $t$ . Assumption 2 implies

$$Q^M(\underline{\theta}_t) < \underline{\theta}_t. \quad (1)$$

The expected value of the deadweight loss in the current period when the regulator uses the *myopic quota*, given the state variable  $\underline{\theta}_t$  is  $M(\underline{\theta}_t)$ :

$$M(\underline{\theta}_t) = E_t [L(Q^M, \theta)].$$

Finally, we define  $P(Q, \underline{\theta}_t)$  as the subjective probability that the quota  $Q$  is binding, given the value of the state variable,  $\underline{\theta}_t$ .

By definition, any quota less than  $\underline{\theta}_t$  is binding with probability 1. That is

$$P(Q, \underline{\theta}_t) = 1 \quad \forall Q \leq \underline{\theta}_t \quad (2)$$

Assumption (3b) implies

$$E_t L(Q, \theta) > E_t L(Q^M(\underline{\theta}_t), \theta), \quad \forall Q > Q^M(\underline{\theta}_t) \quad (3)$$

Equations (2) and (3) imply that a quota in the interval  $(Q^M(\underline{\theta}_t), \underline{\theta}_t]$  creates a loss in welfare in the current period without changing the regulator's information. It is never optimal to use a quota in this interval. If the regulator uses a quota  $Q_t > \underline{\theta}_t$  we say that the regulator "searches in period  $t$ ". The alternative to searching is to use the myopic quota.

A quota in period  $t$ ,  $Q_t$ , that exceeds  $\underline{\theta}_t$  may be slack or binding. In the first case the regulator learns the true value of  $\theta$  and his problem ends. In the second case he knows that  $\theta \geq Q_t$ . In this case,  $\underline{\theta}_{t+1} = Q_t$ . At time  $t$ , given the current quota, the regulator knows what the value of the state variable will be in the next period, conditional upon not having learned the true value of  $\theta$ .

This characteristic holds for an arbitrary number of periods. Given any quota sequence  $\{Q_s\}_{s=t}^T$  the regulator knows that by time  $t'$ , (where  $t < t' \leq T$ ) either one of the previous

quotas will have been slack, or none will have been slack. In the former case he will know the true value of  $\theta$ . In the latter case, the value of  $\underline{\theta}_{t'}$  will be equal to the largest quota between time  $t$  and  $t'$ . This feature makes our control problem fairly simple to solve.

Providing that the future is discounted, it cannot be optimal to use the myopic quota in one period and then search in a subsequent period. Information does not change under the myopic quota. If it is optimal to incur a cost in order to learn, it is better to do it sooner rather than later.

Rather than looking for an optimal policy function that maps the current state into the current control, we can break the problem into two steps. In the first step, the regulator chooses the number of searches, which we denote  $T$ . In the second step, the regulator chooses the optimal conditional quota sequence  $\{Q_s\}_{s=t}^T$ . This quota sequence is conditional in the following sense: the regulator follows it unless one of the quotas has been slack. If one quota is slack, he learns the true value of  $\theta$  and he switches to the first best (full information) quota. Hereafter, when we say that a program involves  $T$  searches, we mean that the regulator intends to search  $T$  more times, conditional upon not learning the true value of  $\theta$  before the  $T$  searches are completed.

Denote the value of the optimal program, i.e. the minimized expectation of the discounted stream of deadweight loss, as  $J(\underline{\theta})$ . Denote  $J^T(\underline{\theta})$  as the value of the optimal program when the regulator decides to search  $T$  times. From the previous comments, it is clear that  $\min_T J^T(\underline{\theta}) = J(\underline{\theta})$ . Our principle result is that it is optimal to search at most one time: the optimal value of  $T$  is either 0 or 1.

We begin by showing that the optimal value of  $T$  is finite. In order to confirm this fact in simple manner, we strengthen Assumption 2 slightly, replacing it with

**Assumption 4**  $Q^M(\underline{\theta}_t) < \underline{\theta}_t - \epsilon$  for some  $\epsilon > 0$ .

Assumption 2, which implies equation (1), does not exclude the possibility that the myopic quota is arbitrarily close to the lower bound of the support,  $\underline{\theta}$ . Assumption 4, on the other hand, states that the myopic quota is bounded away from  $\underline{\theta}$ .

**Lemma 1** *Suppose that Assumption 3 and 4 hold, that  $\bar{\theta} < \infty$ , and that the regulator's subjective distribution of  $\theta$  has finitely many (possibly 0) mass points. Under these assumptions, the optimal  $T$  is finite.*

The appendix contains the proofs of Lemma 1 and the following proposition. The intuition for the lemma is straightforward. Since  $\bar{\theta} < \infty$  an infinite number of searches involves an

infinite number of marginal searches. The cost of a marginal search is non-negligible, but (provided that there is not a mass point at  $\underline{\theta}$ ) the value of information it reveals is negligible. Thus marginal searches are suboptimal and the optimal value of  $T$  is finite.

Our main result is:

**Proposition 1** *Under Assumptions 1 and the other assumptions of Lemma 1, the optimal value of  $T$  is 0 or 1.*

The proof uses the fact that at time  $t$  the regulator can determine the optimal conditional sequence of quotas  $\{Q_s\}_{s=t}^T$ . If the optimal  $T \geq 2$ , at some point the regulator wants to search two more times. Denote  $Q_2$  as the penultimate search when  $T = 2$ , and  $Q_1 > Q_2$  as the final search. The proof that this sequence cannot be optimal demonstrates that the regulator could have done better by skipping the penultimate search ( $Q_2$ ) and immediately setting  $Q_1$ . By following his original plan ( $Q_2$  followed by  $Q_1$ ) one of three things will happen: (1)  $\theta < Q_2$ , where, by Assumption 3b the same loss in welfare is realized by setting  $Q_1$  immediately; (2)  $Q_2 < \theta < Q_1$ , where the loss by setting  $Q_2$  could have been avoided by first setting  $Q_1$ ; or (3)  $Q_2 < Q_1 < \theta$ , where the final lower bound on the support,  $Q_1$  could have been recognized without incurring the loss associated with first setting  $Q_2$ .

We briefly describe several characteristics of the optimal quota. Given our previous assumptions, it is optimal to use the myopic quota when the support of the unknown parameter is small. When the support is small, the amount of uncertainty is small, so the expected dead-weight loss of using the myopic quota is small. The cost of a search may nevertheless be large, since it might be necessary to use a large quota in order for there to be a possibility that it is not binding. Without stronger assumptions, we cannot rule out the possibility that for some values of  $\underline{\theta}$  it is optimal to search, and that there exist both larger and smaller values of  $\underline{\theta}$  at which it is optimal not to search.

When the regulator ignores the future ( $\beta = 0$ ) the myopic quota is obviously optimal. By continuity, the myopic quota is optimal for small values of  $\beta$ . When  $\beta = 1$  the optimal quota is  $\bar{\theta}$ , provided that the loss under this quota is finite. By continuity, it is optimal to use the quota  $\bar{\theta}$  for values of  $\beta$  close to 1.<sup>3</sup> More generally, the optimal quota is a nondecreasing function of  $\beta$ .

---

<sup>3</sup>The case where  $\beta \approx 1$  and the case where the length of a period is extremely small are equivalent. We mentioned above that if the length of the period is small, it is optimal to incur (possibly large) costs for a short period of time in order to avoid future costs.

### 3 Conclusion

A firm's response to an emissions tax tells the regulator something about the firm's abatement cost curve. If a quota is binding with probability 1, and quota rights are not traded, the quota is not informative about costs. This informational difference between taxes and quotas has not (to our knowledge) previously been explored.

A quota can be used to discover information about abatement costs if there is a positive probability that it is not binding, and if the firm's unregulated level of emissions is correlated with its abatement costs. We have shown how quotas can be used to learn in an environment where the marginal cost and marginal damage curves are constant and firms are non-strategic.

In this situation, the regulator faces what appears to be a standard search problem, similar to the problem of increasing capacity to discover the size of a market. Two features of the search problem with quotas alter the characteristics of the optimal search. There is a non-negligible cost of acquiring even a small amount of information, but the cost of using a quota that is slack does not increase as the quota increases. These features mean that gradual search is never optimal. The regulator may or may not search in the first period, but whatever the outcome in that period, he does not search again.

## 4 Appendix: Proofs

**Proof.** (Lemma 1) Define

$$F(\underline{\theta}_t) \equiv E_t L(\underline{\theta}_t, \theta) - M(\underline{\theta}_t).$$

$F(\underline{\theta}_t)$  is the expected additional cost of using the quota  $\underline{\theta}_t$  rather than the myopic quota. Only quotas greater than  $\underline{\theta}_t$  have a positive probability of being slack. The regulator learns only by using a quota greater than  $\underline{\theta}_t$ , so we can regard  $F(\underline{\theta}_t)$  as a “fixed cost” of learning. Under assumptions 3 and 4,  $F(\underline{\theta}_t) \geq \delta > 0$  for some number  $\delta$ .

Suppose that at time  $t$ , given  $\underline{\theta}_t$ , the regulator intends to “search”  $T$  times (i.e., he intends to use quotas greater than the contemporaneous myopic quota  $T$  times, conditional on not having yet learned the value of  $\theta$ ). In this case, the average, over these  $T$  searches, of the difference between the quota and the contemporaneous lower bound of the support, is no greater than  $\frac{\bar{\theta} - \underline{\theta}_t}{T}$ . (If this were not true, then by the time of the  $T$ ’th search the quota is greater than the upper bound  $\bar{\theta}$ .) If  $T = \infty$ , the regulator plans to use infinitely many quotas that are arbitrarily close to the contemporaneous lower bound. We denote such a quota as a “marginal search”.

In order to show that  $T = \infty$  is not optimal, it is sufficient to show that the program associated with  $T = \infty$  involves at least one action which is not optimal given the current state. (The existence of such an action violates the Principle of Optimality and therefore cannot be part of an optimal program.)

The current cost of a marginal search is no less than  $\delta > 0$ . If the subjective distribution does not have a mass point at  $\underline{\theta}_t$  the information provided by a marginal search is negligible. In that case, the cost of the search must exceed the value of the information it provides, so such a search cannot be optimal. Since (by assumption) there are a finite number of mass points of the subjective distribution, the program with  $T = \infty$  involves infinitely many marginal searches where the cost is strictly positive and the benefit is negligible. ■

**Proof.** (Proposition 1) By Lemma 1 we need only consider finite values of  $T$ . Suppose, contrary to the Proposition, that the optimal value is  $T \geq 2$ . In this case, there exists a value of  $\underline{\theta}$  at which it is optimal to search exactly two more times (conditional on not first learning the true value of  $\theta$ ). Thus, it is sufficient to show that it cannot be optimal to search two more times.

Suppose to the contrary that for some value of  $\underline{\theta}$  it is optimal to search 2 more times. Using previous notation, and letting  $\beta$  equal the discount factor, the dynamic programming equation

for this problem is

$$J(\underline{\theta}) = J^2(\underline{\theta}) = \min_Q [E_2 \{L(Q, \underline{\theta}) + \beta \{P(Q, \underline{\theta})J^1(Q) + [1 - P(Q, \underline{\theta})]0\}] \quad (4)$$

$$= E_2 \{L(Q_2, \underline{\theta}) + \beta \{P(Q_2, \underline{\theta})J^1(Q_2)\}. \quad (5)$$

We use  $E_i$ ,  $i = 1, 2$ , to denote the expectation over the unknown parameter  $\theta$ , conditioned on beliefs at the time when it is optimal to search (at most)  $i$  more times;  $Q_i$  is the optimal quota when there are (at most)  $i$  remaining searches.

If the quota is not binding (which occurs with probability  $1 - P(Q, \underline{\theta})$ ), the regulator learns the true value of  $\theta$ , and future losses are 0, by Assumption (3c). If the quota is binding, the value of the state in the next period is  $Q$ . By the hypothesis that we seek to falsify, it is optimal to search in the next period if the regulator has not learned in the true value of  $\theta$ .  $J^1(Q)$  is the optimal value of the program when it is optimal to search one more time, and the current state is  $Q$ .

At the time of the penultimate search, when the regulator uses  $Q_2$ , he knows that  $\theta$  is either greater or less than  $Q_2$ . Thus, we can write

$$E_2 \{L(Q_2, \underline{\theta})\} = [E_{\theta \geq Q_2} \{L(Q_2, \underline{\theta})\} \{P(Q_2, \underline{\theta})\}] + [E_{\theta \leq \theta < Q_2} \{L(Q_2, \underline{\theta})\} \{1 - P(Q_2, \underline{\theta})\}]. \quad (6)$$

The expectations on the right side of equation (6) are conditioned on all of the regulator's information at the time of the penultimate search, in addition to the information contained in the inequalities in the subscript of  $E$ . Thus, for example,  $E_{\theta \geq Q_2}(\cdot)$  is an abbreviation for  $E_{\theta \geq Q_2}(E_2(\cdot))$ .

Using this expression we can rewrite equation (5) as

$$J^2(\underline{\theta}) = [E_{\theta \geq Q_2} \{L(Q_2, \underline{\theta}) + \beta J^1(Q_2)\} \{P(Q_2, \underline{\theta})\}] + [E_{\theta \leq \theta < Q_2} \{L(Q_2, \underline{\theta})\} \{1 - P(Q_2, \underline{\theta})\}]. \quad (7)$$

The value function in the next period – the final searching period – (assuming that the regulator has not learned the value of  $\theta$ ) is given by

$$J^1(Q_2) = \min_Q \left[ E_1 \{L(Q, \underline{\theta})\} + \frac{\beta}{1 - \beta} M(Q) P(Q, Q_2) \right] \quad (8)$$

$$= E_1 \{L(Q_1, \underline{\theta})\} + \frac{\beta}{1 - \beta} M(Q_1) P(Q_1, Q_2) \quad (9)$$

Note that the unknown parameter  $\theta$  appears in only the first term on the right side of equation (9).



Consider an alternative program that involves a single search using  $Q_1$  (rather than  $Q_2$ ) in the “first period”, when the state is  $\underline{\theta}$ . Denote the expected value of this program as  $A(Q_1, \underline{\theta})$ . We write this expected value by conditioning on the two events:  $\theta \geq Q_2$  and  $\theta < Q_2$ :

$$A(Q_1, \underline{\theta}) = \left[ E_{\theta \geq Q_2} \{L(Q_1, \theta)\} + \frac{\beta}{1-\beta} M(Q_1) P(Q_1, Q_2) \right] P(Q_2, \underline{\theta}) + [E_{\theta \leq \theta < Q_2} \{L(Q_1, \theta)\} \{1 - P(Q_2, \underline{\theta})\}] \quad (10)$$

Assumption (3b) implies

$$L(Q, \theta) = L(Q', \theta), \quad \forall Q, Q' \geq \theta$$

This equation states that if two quotas are slack, they lead to the same loss in current welfare. The quota used in the last search,  $Q_1$ , involves learning. Therefore  $Q_1 > Q_2$ . Consequently, we have

$$E_{\theta \leq \theta < Q_2} \{L(Q_2, \theta)\} = E_{\theta \leq \theta < Q_2} \{L(Q_1, \theta)\}.$$

Note that the left and the right side of this equation are, respectively, equal to the last term in square brackets in equation (7) and the last term in square brackets in equation (10). In addition we see (using equation (9)) that the first term in square brackets in equation (10) is equal to  $J^1(Q_2)$ .

Thus, we can write  $A(Q_1, \underline{\theta})$  as

$$A(Q_1, \underline{\theta}) = J^1(Q_2) P(Q_2, \underline{\theta}) + [E_{\theta \leq \theta < Q_2} \{L(Q_2, \theta)\} \{1 - P(Q_2, \underline{\theta})\}] \quad (11)$$

Using equations (7) and (11) we write the difference in payoffs as

$$J^2(\underline{\theta}) - A(Q_1, \underline{\theta}) = [E_{\theta \geq Q_2} \{L(Q_2, \theta)\} + (\beta - 1) J^1(Q_2)] \{P(Q_2, \underline{\theta})\}. \quad (12)$$

If  $P(Q_2, \underline{\theta}) = 0$  the regulator learns the true value of  $\theta$  with probability 1 during the first search. However, by assumption, the regulator intends to search a second time. Therefore  $P(Q_2, \underline{\theta}) > 0$ . Consequently, to obtain a contradiction we need to show that the term in square brackets in equation (12) is positive. This positive value means that expected discounted stream of costs are higher under the optimal two-search strategy than under an alternative. The two-search strategy can therefore not be optimal.

In other words, we must show

$$\frac{E_{\theta \geq Q_2} \{L(Q_2, \theta)\}}{1 - \beta} > J^1(Q_2). \quad (13)$$

Note that  $E_{\theta \geq Q_2}(\cdot) = E_{\theta \geq Q_2}(E_2(\cdot)) = E_1(\cdot)$ . The first equality is merely a restatement of an earlier abbreviation, and the second equality states that the only additional information that the regulator has at the time of the last search, that he did not have at the time of the penultimate search, is that  $\theta \geq Q_2$ . Therefore we can rewrite equation (13) as

$$\frac{E_1\{L(Q_2, \theta)\}}{1 - \beta} > J^1(Q_2) \quad (14)$$

The left side of inequality (14) is the present discounted value of using, in perpetuity, the previous quota (which the regulator knows will be binding). However, the costs of that policy exceed the cost of the optimal policy, which requires searching one more time. Consequently inequality (14) is satisfied, and we have a contradiction. ■

## References

- HOEL, M., AND L. KARP (1998): "Taxes Versus Quotas for a Stock Pollutant," *Fondazione Eni Enrico Mattei, Nota di Lavoro*, 29.98, <http://are.Berkeley.EDU/karp/>.
- HOEL, M., AND L. KARP (2000): "Taxes and Quotas for a Stock Pollutant with Multiplicative Uncertainty," *Journal of Public Economics*, forthcoming.
- KARP, L., AND J. ZHANG (1999): "Regulation of Stock Externalities with Learning," Working Paper No. 892, Department of Agricultural and Resource Economics, University of California, Berkeley; <http://are.Berkeley.EDU/karp/>.
- (2000): "Stock Externality Regulation with Endogenous Technical Change and Learning," Department of Agricultural and Resource Economics, University of California, Berkeley; <http://are.Berkeley.EDU/karp/>.
- MALCOMSON, J. (1978): "Prices vs. Quantities: A Critical Note on the Use of Approximations," *Review of Economic Studies*, 45(1), 203–07.
- NEWELL, R., AND W. PIZER (1998): "Stock Externality Regulation Under Uncertainty," Resources for the Future Working Paper.
- ROB, R. (1991): "Learning and Capacity Expansion Under Demand Uncertainty," *Review of Economic Studies*, 58, 655–675.
- ROBERTS, M., AND A. SPENCE (1976): "Effluent charges and licenses under uncertainty," *Journal of Public Economics*, 5(3-4), 193–208.
- STAVINS, R. (1996): "Correlated uncertainty and policy instrument choice," *Journal of Environmental Economics and Management*, 30(2), 218–232.
- WATSON, W., AND R. RIDKER (1984): "Losses from effluent taxes and quotas under uncertainty," *Journal of Environmental Economics and Management*, 11(4), 310–26.
- WEITZMAN, M. (1974): "Prices versus quantities," *Review of Economic Studies*, 41(4), 477–91.
- YOHE, G. (1977): "Comparisons of price and quantity controls: A survey," *Journal of Comparative Economics*, 1(3), 213–33.