

## OPTIMAL RESEARCH AND DEVELOPMENT EXPENDITURES UNDER AN INCREMENTAL TAX INCENTIVE SCHEME\*

P. L'ÉCUYER

*Département d'informatique, Université Laval, Sainte-Foy, Canada, G1K 7P4*

A. HAURIE

*GERAD, École des Hautes Études Commerciales, 5255 Decelles, Montréal, H3T 1V6, Canada*

A. HOLLANDER

*Département d'économie, Université de Montréal, Case Postale 6128, Succursale « A », Montréal, H3C 3J7, Canada*

Received August 1984

Revised February 1985

The effect of an incremental tax incentive on optimal R and D spending of a profit maximizing firm is investigated. A dynamic programming approach permits the complete characterization of an optimal policy for a one-period-memory piecewise linear incentive scheme, and efficient computation of an  $\epsilon$ -optimal policy in the more general case.

dynamic programming \* economics \* investment \* taxation

### 1. Introduction

In an effort to stimulate innovation in the U.S. economy, the Economic Recovery Act instituted a tax credit for incremental R and D spending. It provides for a 25% tax credit on R and D expenditures (RDE) incurred between June 30, 1981 and December 31, 1985, which exceed the average RDE during the three previous fiscal years. A similar provision, enacted in Canada in 1977, remained in force until 1983.

The purpose of this paper is to investigate the effect of such incremental incentive schemes on the RDE pattern of a profit maximizing firm. The fundamental result, from a practical point of view, is that such incentives may induce a cyclical pattern in the RDE. This is basically due to the conjunction of a hereditary dynamical structure

and a nonconcave pay-off function. In Section 2 the determination of an optimal RDE policy under incentive is posed as a deterministic optimal control problem in a discrete time setting. A solution to the problem is looked for, using recent developments in dynamic programming. Section 3 studies in depth a piecewise linear incentive scheme with a one period memory. The cyclical pattern of optimal RDE is established and a surrogate static optimization problem is proposed for the computation of the optimal policy. The more general problem corresponding to the actual incentives (3 period memory) is approached numerically in Section 4. A successive approximation algorithm, with approximate computation of the transform of the value function by the dynamic programming operator, is implemented. The pattern of an optimal RDE policy is then exhibited, and an optimal cyclical RDE is obtained.

For the sake of simplicity of exposition, it has been assumed that except for the incentive, RDE has an immediate effect on profits. The model could be straightforwardly extended to the case

\* Research of the first author supported by NSERC grant A5463 and of the second author by SSHRC, grant 410-81-0722, FCAC grant 0428, NSERC grant A4952.

where there is a multi-period effect of RDE in addition to that of the incentive.

**2. Optimal RDE under incentive: A dynamic programming approach**

The firm we consider picks a level  $a_t$  of RDE within a compact interval  $A \triangleq [0, b]$ , at each period  $t$  belonging to an infinite sequence  $T \triangleq \{0, 1, 2, \dots\}$  of periods. We assume that the net expected profit associated with  $a_t$  is concentrated in period  $t$  and given by a bounded upper semi-continuous real function  $f: A \rightarrow \mathbb{R}$ .

In the absence of incentive the firm seeks to maximize the net discounted sum of profits ( $\alpha \in (0, 1)$  is the discount factor):

$$\sum_{t=0}^{\infty} \alpha^t f(a_t). \tag{1}$$

The obvious optimal RDE sequence is given by  $a_t \equiv \bar{a}$  with  $\bar{a} = \arg \max\{f(a): a \in A\}$ .

Consider now the effects of a tax incentive scheme which is a function of present RDE  $a_t$  and the string  $x_t \triangleq (a_{t-m}, a_{t-m+1}, \dots, a_{t-1})$  or RDE chosen in the past  $m$  periods. The incentive scheme is defined as a bounded upper semi-continuous (u.s.c.) function  $h: X \times A \rightarrow \mathbb{R}$  where  $X \triangleq A^m$ .

Given the initial string  $x_0$ , the firm seeks to maximize

$$\begin{aligned} \Phi(\{a_t\}_{t \in T}, x_0) &= \sum_{t=0}^{\infty} \alpha^t g(x_t, a_t) \\ &\triangleq \sum_{t=0}^{\infty} \alpha^t (f(a_t) + h(x_t, a_t)) \end{aligned} \tag{2}$$

subject to

$$x_t = (a_{t-m}, a_{t-m+1}, \dots, a_{t-1}), \quad t \in T, \tag{3}$$

$$x_0 = (a_{-m}, a_{-m+1}, \dots, a_{-1}) \in X \text{ given.} \tag{4}$$

Clearly the incentive scheme complicates the optimization which is now akin to a discrete time optimal control problem, with  $x_t$  and  $a_t$  being respectively the state and control variables at  $t$ . We shall review the dynamic programming approach for the solution of this problem.

Let constants  $g_0$  and  $g_1$  be respectively lower and upper bounds for  $g(x, a)$ ,  $x \in X$ ,  $a \in A$  (such bounds exist), and let  $B$  be the class of u.s.c. functions

$$V: X \rightarrow [g_0/(1 - \alpha), g_1/(1 - \alpha)].$$

We define, for any  $V$  in  $B$ ,

$$H(V)(x, a) \triangleq g(x, a) + \alpha V(a_{-m+1}, \dots, a_{-1}, a),$$

$$x = (a_{-m}, \dots, a_{-1}) \in X, \quad a \in A, \tag{5}$$

$$T(V)(x) = \sup_{a \in A} H(V)(x, a), \quad x \in X, \tag{6}$$

and for any function  $\mu: X \rightarrow A$ , called a policy,

$$T_\mu(V)(x) = H(V)(x, \mu(x)), \quad x \in X. \tag{7}$$

Since  $A$  is compact and  $V$  is u.s.c., the supremum is always attained in (6). The operators  $T_\mu: B \rightarrow B$  and  $T: B \rightarrow B$  are the usual dynamic programming operators. We now can state the following well known results.

**Proposition 1.** For any function  $\mu: X \rightarrow A$ , define

$$V_\mu(x_0) = \sum_{t=0}^{\infty} \alpha^t g(x_t, a_t) \tag{8}$$

s.t. (3), (4) and

$$a_t = \mu(x_t), \quad t \in T. \tag{9}$$

Also define

$$V_*(x_0) = \sup_{\mu} V_\mu(x_0). \tag{10}$$

Then the following hold:

(i)  $V_*$  is in  $B$  and is the unique solution to

$$T(V) = V, \quad V \in B. \tag{11}$$

(ii)  $V_\mu$  is in  $B$  and is the unique solution to

$$T_\mu(V) = V, \quad V \in B. \tag{12}$$

(iii) If  $T^n$  denotes the  $n$ -fold composition of  $T$ ,

$$\lim_{n \rightarrow \infty} \sup_{x \in X} |T^n(V)(x) - V_*(x)| = 0 \quad \forall V \in B. \tag{13}$$

(iv)  $V_\mu = V_*$  iff  $T_\mu(V_*) = V_*$  and iff  $T(V_\mu) = V_\mu \in B$ .

(v) For a given  $V$  in  $B$ , there exists a sequence of policies  $\mu_1, \mu_2, \dots$  such that

$$T^n(V) = T_{\mu_n}(T^{n-1}(V)), \quad n \in \mathbb{N}, \tag{14}$$

and

$$\lim_{n \rightarrow \infty} \sup_{x \in X} |V_{\mu_n}(x) - V_*(x)| = 0.$$

**Proof:** see [1-6].

**3. Optimal RDE under a piecewise linear incentive with one period-memory**

We characterize in this section an optimal RDE policy when the incentive scheme has a one-period memory ( $m = 1, x_t = a_{t-1}$ ) and is defined by

$$h(x, a) = \delta \min(0, a - x) + \gamma \max(0, a - x). \tag{15}$$

With parameters  $\delta$  and  $\gamma$  representing respectively the penalty per unit of RDE decrease below the previous level and the premium (or tax credit) per unit of RDE increase above the previous level.

A related model was studied in [8], but with the important restriction  $\gamma = \delta$ , under which  $h$  is linear.

According to (5) and (6) we have

$$H(V)(x, a) = \begin{cases} f(a) + \alpha V(a) + \delta(a - x) & \text{if } a \leq x, \\ f(a) + \alpha V(a) + \gamma(a - x) & \text{if } a \geq x. \end{cases} \tag{16}$$

$$T(V)(x) = \max \left\{ \max_{0 \leq a \leq x} (f(a) + \alpha V(a) + \delta a) - \delta x, \max_{x \leq a \leq b} (f(a) + \alpha V(a) + \gamma a) - \gamma x \right\}. \tag{17}$$

The two propositions that follow state that the optimal policy dictates a constant RDE sequence when  $\delta \geq \gamma$  and that, when  $\delta < \gamma$ , the optimal RDE sequence is either constant or periodic with period 2. Notice that no special assumption is made on  $f$  other than boundedness and upper semi-continuity.

**Proposition 2.** *If  $\delta \geq \gamma$ , then a constant RDE sequence  $a_t \equiv a_*$ ,  $\forall t \in T$  ( $a_*$  may depend on  $x_0$ ) maximizes (2) subject to (3, 4, 15), and*

$$V_*(x) = \max \left\{ \max_{0 \leq a \leq x} (\delta(a - x) + f(a)/(1 - \alpha)), \max_{x \leq a \leq b} (\gamma(a - x) + f(a)/(1 - \alpha)) \right\}. \tag{18}$$

**Proof:** By Proposition 1(i) we know that  $V_*$  exists.

For a given  $x_0$  in  $X$ , let

$$a_* = \arg \max_{a \in A} H(V_*)(x_0, a).$$

Two cases are possible: If  $a_* \leq x_0$ , then

$$f(a_*) + \alpha V_*(a_*) + \delta(a_* - x_0) \geq f(a) + \alpha V_*(a) + \begin{cases} \delta(a - x_0) & \text{if } 0 \leq a \leq x_0, \\ \gamma(a - x_0) & \text{if } x_0 \leq a \leq b, \end{cases}$$

from which we obtain

$$f(a_*) + \alpha V_*(a_*) \geq \begin{cases} f(a) + \alpha V_*(a) + \delta(a - a_*) & \text{if } a \leq a_*, \\ f(a) + \alpha V_*(a) + \delta(a - a_*) \geq f(a) + \alpha V_*(a) + \gamma(a - a_*) & \text{if } a_* \leq a \leq x_0, \\ f(a) + \alpha V_*(a) + \gamma(a - x_0) + \delta(x_0 - a_*) \geq f(a) + \alpha V_*(a) + \gamma(a - a_*) & \text{if } x_0 \leq a. \end{cases} \tag{19}$$

If  $a_* > x_0$ , then

$$f(a_*) + \alpha V_*(a_*) + \gamma(a_* - x_0) \geq f(a) + \alpha V_*(a) + \begin{cases} \delta(a - x_0) & \text{if } 0 \leq a \leq x_0, \\ \gamma(a - x_0) & \text{if } x_0 \leq a \leq b, \end{cases}$$

and

$$f(a_*) + \alpha V_*(a_*) \geq \begin{cases} f(a) + \alpha V_*(a) + \delta(a - x_0) - \gamma(a_* - x_0) \geq f(a) + \alpha V_*(a) + \delta(a - a_*) & \text{if } a \leq x_0, \\ f(a) + \alpha V_*(a) + \gamma(a - a_*) \geq f(a) + \alpha V_*(a) + \delta(a - a_*) & \text{if } x_0 \leq a \leq a_*, \\ f(a) + \alpha V_*(a) + \gamma(a - a_*) & \text{if } a_* \leq a. \end{cases} \tag{20}$$

Since for  $x = a = a_*$ ,  $g(a_*, a_*) = f(a_*)$ , (19) and (20) establish that

$$V_*(a_*) = f(a_*) = \alpha V_*(a_*).$$

Therefore, by Proposition 1(ii), (iv), the optimal policy calls for  $a_0 = \mu(x_0) = a_*$ ,  $a_t = \mu(a_*) = a_*$ ,  $\forall t \geq 1$ . Notice that the optimal investment sequence, albeit constant, depends on the initial state  $x_0$ .

Since the RDE level  $a_0$  chosen at  $t = 0$  is maintained forever, we have  $h(x_t, a_t) = 0$  for all  $t > 0$ , and the total discounted profit is thus given by one of the two expressions in the RHS of (18).

**Proposition 3.** *If  $\delta < \gamma$  in (15), then there exist three constants  $a_*$ ,  $a_{**}$  and  $x_*$  in  $A$ , with  $a_* \leq x_* \leq a_{**}$ , such that the optimal RDE policy  $\mu_*$  is defined by*

$$\mu_*(x) = \begin{cases} a_* & \text{if } x > x_*, \\ a_{**} & \text{if } x \leq x_*, \end{cases} \quad (21)$$

where

$$x_* \triangleq (v_* - u_*) / (\gamma - \delta), \quad (22)$$

$$v_* \triangleq f(a_{**}) + \alpha V_*(a_{**}) + \gamma a_{**}, \quad (23)$$

$$u_* \triangleq f(a_*) + \alpha V_*(a_*) + \delta a_*. \quad (24)$$

**Proof.** In that case (17) can be rewritten as

$$T(V)(x) = \max(u - \delta x, v - \gamma x)$$

with

$$u = \max_{0 \leq a \leq x} (f(a) + \alpha V(a) + \delta a), \quad (25)$$

$$v = \max_{x \leq a \leq b} (f(a) + \alpha V(a) + \gamma a). \quad (26)$$

When  $V = V_*$ , we denote by  $u_*$  and  $v_*$  the respective values of  $u$  and  $v$ , and by  $a_*$  and  $a_{**}$  the largest maximizing values of  $a$  for a given value of  $x$  in (25) and (26) respectively. An optimal RDE policy is then to spend  $a_*$  if  $u_* - \delta x > v_* - \gamma x$ , and  $a_{**}$  otherwise. This leads to (21–22).

It remains to prove that  $a_* \leq x_* \leq a_{**}$ . From (22–24) and the definition of  $a_*$  and  $a_{**}$ , we have

$$\begin{aligned} & (\gamma - \delta)(x_* - a_*) \\ &= v_* - u_* - \gamma a_* + \delta a_* \\ &= f(a_{**}) + \alpha V_*(a_{**}) + \gamma a_{**} - f(a_*) \\ &\quad - \alpha V_*(a_*) - \gamma a_* \\ &\geq 0, \end{aligned}$$

and

$$\begin{aligned} & (\gamma - \delta)(a_{**} - x_*) = u_* - v_* + \gamma a_{**} - \delta a_{**} \\ &= f(a_*) + \alpha V_*(a_*) + \delta a_* - f(a_{**}) - \alpha V_*(a_{**}) \\ &\quad - \delta a_{**} \geq 0. \end{aligned}$$

This establishes the desired inequalities. Notice that the optimal RDE sequence is constant when  $a_{**} = x_*$ .

The following corollary provides a simple mean to compute the optimal policy when  $\delta < \gamma$ .

**Corollary 1.** *Consider the constrained (static) optimization problem*

$$\text{Max } u + v \quad (27)$$

s.t.

$$u - \alpha v = f(a_1) + (\delta - \alpha\gamma)a_1, \quad (28)$$

$$v - \alpha u = f(a_2) + (\gamma - \alpha\delta)a_2, \quad (29)$$

$$v - u \geq (\gamma - \delta)a_1, \quad (30)$$

$$v - u \leq (\gamma - \delta)a_2, \quad (31)$$

$$a_1, a_2 \in A,$$

$$u, v \in \mathbb{R}.$$

Then the optimal policy defined by (21–24) provides a solution to (27–31) when  $u = u_*$ ,  $v = v_*$ ,  $a_1 = a_*$ ,  $a_2 = a_{**}$ . Conversely, if  $(a_0, a_{00}, u_0, v_0)$  is a solution to (27–31), then the policy defined by

$$\mu_0(x) = \begin{cases} a_{00} & \text{if } x > \bar{x}, \\ a_{00} & \text{if } x \leq \bar{x}, \end{cases} \quad (32)$$

$$\bar{x} = (v_0 - u_0) / (\gamma - \delta) \quad (33)$$

is optimal.

**Proof.** The function  $V_*$  associated with  $\mu_*$  defined by (21–24) is given by

$$V_*(x) = \begin{cases} u_* - \delta x & \text{for } x > x_*, \\ v_* - \gamma x & \text{for } x \leq x_*. \end{cases} \quad (34)$$

From (23, 24, 34), we obtain (28, 29). The inequalities (30, 31) are also implied by  $a_* \leq x_* \leq a_{**}$  and (22). Therefore  $(u_*, v_*, a_*, a_{**})$  is feasible for the constrained optimization problem (27–31).

Let  $(u_0, v_0, a_0, a_{00})$  be any other feasible solution to (28–31). Define a policy  $\mu_0$  as in (32–33). By straightforward calculation, one obtains

$$T_{\mu_0}(V_0)(x) = V_0(x) \quad \forall x \in A \quad (35)$$

where

$$V_0(x) = \begin{cases} u_0 - \delta x & \text{if } \bar{x} < x \leq b, \\ v_0 - \gamma x & \text{if } 0 \leq x \leq \bar{x}. \end{cases} \quad (36)$$

Therefore  $V_0 = V_{\mu_0}$  by Proposition 1(i), and thus by construction  $V_*(x) \geq V_0(x) \quad \forall x \in A$ . If  $x_* < b$  then

$$u_* = V_*(b) + \delta b \geq V_0(b) + \delta b \geq u_0. \quad (37)$$

If  $x_* = b$ , then necessarily  $a_* = x_* = a_{**} = b$  and

$$\begin{aligned} u_* &= f(b) + \alpha V_*(b) + \delta b = v_* - (\gamma - \delta)b \\ &= V_*(b) + \delta b \geq V_0(b) + \delta b \geq u_0. \end{aligned} \quad (38)$$

Finally

$$v_* = V_*(0) \geq V_0(0) = v_0. \quad (39)$$

From (37–39) we conclude that  $u_* + v_* \geq u_0 + v_0$ .

Further, if  $(u_0, v_0, a_0, a_{00})$  is a solution to the optimization problem (27–31), and if the policy  $\mu_0$  defined by (32–33) is non optimal, then for some  $x$  in  $[0, b]$ ,  $V_0(x) < V_*(x)$  and by (34–36), either  $u_0 < u_*$  or  $v_0 < v_*$ . Since, by assumption,  $u_0 + v_0 \geq u_* + v_*$ , one has either  $u_0 > u_*$ , or  $v_0 > v_*$ , which contradicts (23, 24). Therefore  $\mu_0$  is an optimal policy.

**4. A computational approach for the general case**

The U.S. incentive program, as well as the Canadian one effective till 1983, uses a scheme with a 3 period memory ( $m = 3$ ) and a function  $h(x, a)$  defined by

$$h(x, a) = \gamma \max(0, a - (x_1 + x_2 + x_3)/3) \quad (40)$$

where  $x = (x_1, x_2, x_3)$ .

Even for such a simple piecewise linear incentive scheme, the only possible approach for determining an optimal RDE policy seems to be a direct numerical solution of the dynamic equation. The approach proposed in this section is a successive approximation method based on Proposition 1(v) but with an approximate computation of  $T(V)$  at each iteration. The algorithm proceeds as follows.

*Step 1:* Set  $\underline{V} = 0, \bar{V} = g_1/(1 - \alpha)$ , choose any  $V$  in  $B$  and  $\epsilon > 0$ .

*Step 2:* Compute  $T(V)$  at a finite number of points in  $X$ . Obtain a function  $W$  in  $B$  which approximates  $T(V)$  on  $X$  (e.g. use finite element interpolation). This step can be repeated, resetting  $V := W$ , if one wishes.

*Step 3:* Compute or estimate  $d^- \geq 0, d^+ \geq 0, c^- \geq 0, c^+ \geq 0$  s.t.

$$-d^- \leq T(V) - W \leq d^+, \quad (41)$$

$$-c^- \leq W - V \leq c^+. \quad (42)$$

Redefine

$$\underline{V} := \max(\underline{V}, W - (d^- + \alpha c^-)/(1 - \alpha)), \quad (43)$$

$$\bar{V} := \min(\bar{V}, W + (d^+ + \alpha c^+)/(1 - \alpha)). \quad (44)$$

*Step 4:* Find a policy  $\mu$  and  $\epsilon_0 > 0$  s.t.

$$T_\mu(V) \leq W + \epsilon_0. \quad (45)$$

*Step 5:* Compute  $\epsilon_1 = (\epsilon_0 + d^- + \alpha(c^- + c^+))/(1 - \alpha)$ . If  $\epsilon_1$  is smaller than  $\epsilon$ , then stop. Otherwise return to step 2.

**Proposition 4.** (a) After each iteration (steps 2 to 5),  $\underline{V}$  and  $\bar{V}$  satisfy

$$\underline{V}(x) \leq V_*(x) \leq \bar{V}(x) \quad \forall x \in X. \quad (46)$$

(b) The policy  $\mu$  obtained in step 4 is  $\epsilon_1$ -optimal, i.e.

$$V_\mu(x) > V_*(x) - \epsilon_1 \quad \forall x \in X. \quad (47)$$

(c) If the sequences of values  $d^-, d^+, \epsilon_0$  obtained at successive iterations converge to 0, then  $\underline{V} - \bar{V}$  converges to 0, and, for any  $\epsilon \geq 0$ , an  $\epsilon$ -optimal policy is obtained in a finite number of iterations.

**Proof.** It follows from Theorem 3.3 in [7].

An interactive FORTRAN program was built to implement this algorithm, and used to solve the following numerical example.

**5. A numerical illustration**

Let  $\gamma = 0.5, \alpha = 0.9, b = 4$  and

$$f(a) = 2 \ln(1 + a) - a.$$

At each iteration, we partition  $X$  into rectangular boxes, evaluate  $T(V)$  at the corner points of the boxes, and define  $W$  as the boxwise trilinear interpolating function. For instance, choosing a regular mesh of 5 points along each axis,  $X$  can be partitioned into  $4^3 = 64$  identical cubes. After 60 iterations with this partition, one may refine the grid to obtain  $9^3$  cubic boxes, perform another 20 iterations, refine again into 15 boxes and perform 5 more iterations. One then obtains  $c^- = 0.00088$ ,

Table 1

PERIOD	RDE AMOUNT
-3	1.000
-2	1.000
-1	1.000
0	2.076
1	3.000
2	0.784
3	0.593
4	0.558
5	2.078
6	3.000
7	0.784
8	0.593
9	0.558
10	2.078

$c^+ = 0.00000$ ,  $\bar{V} - \underline{V} = 0.017$ ,  $V(1, 1, 1) = 5.34$  and  $\epsilon_1 \cong 0.10$ . The values of  $d^-$  and  $d^+$  have been estimated by recomputing  $T(V)$  on a much finer grid, and comparing these values of  $T(V)$  with the actual values of  $W$ . The retained policy  $\mu$  is defined implicitly by

$$\mu(x) = \arg \max_{a \in A} H(W)(x, a).$$

Using this policy, it is easy to compute an investment schedule. If e.g. the RDE amounts for the last 3 periods have been chosen to maximize  $f(a)$ , i.e.  $\bar{a} = 1$ , we obtain Table 1.

From period 2, this schedule exhibits a five period cycle for which  $h(x, a)$  is positive for the last 2 periods of the cycle.

On the other hand, if one assumes a cyclic schedule of period 5 where the firm takes advantage of the incentive only during the last 2 periods, the total present value of the schedule at the beginning of a cycle is

$$\begin{aligned} J(a_1, a_2, a_3, a_4, a_5) &= \{ f(a_1) + \alpha f(a_2) + \alpha^2 f(a_3) \\ &+ \alpha^3 (f(a_4) + \gamma(a_4 - (a_1 + a_2 + a_3)/3)) \\ &+ \alpha^4 (f(a_5) + \gamma(a_5 - (a_2 + a_3 + a_4)/3)) \} \\ &\times (1 - \alpha^5)^{1/2}. \end{aligned}$$

This function of five variables is maximized for  $a_1 = 0.7833$ ,  $a_2 = 0.5917$ ,  $a_3 = 0.5564$ ,  $a_4 = 2.0769$  and  $a_5 = 3.0$ . This is in very close agreement with the approximation obtained through the D.P. algorithm.

## 6. Conclusion

It is interesting to notice that the optimal periodicity is 5 although the memory of the system is over the last 3 periods.

The cyclical pattern of R and D expenditures induced by an incremental tax incentive scheme should lead to further investigation of the following topics: (1) the optimal design of governmental incentive programs for which a balance has to be found between the cost of the program and its pay-off; (2) extension of the approach to a continuous-time optimal control setting. The first question pertains to economics, the second is related to the theory of periodic control, both topics pose interesting challenges as nonstandard optimization problems.

## References

- [1] D. Blackwell, "Discounted dynamic programming", *Ann. Math. Statist.* 36, 226-235 (1965).
- [2] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, 1978.
- [3] A. Maitra, "Discounted dynamic programming on compact metric spaces", *Sankhya: The Indian Journal of Statistics: Series A*, 30A, 211-216 (1968).
- [4] K. Hinderer, *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter*, Lect. Notes in Oper. Res. and Math. Syst. 33, Springer-Verlag, 1970.
- [5] P. L'Écuyer, "Processus de décision Markoviens à étapes discrètes: Application à des problèmes de remplacement d'équipement", les cahiers du GERAD, no. G-83-06, École des H.E.C., Montréal, 1983.
- [6] P. L'Écuyer and A. Haurie, "Discrete event dynamic programming in Borel spaces with state dependant discounting", report no. DIUL-RR-8309, Département d'informatique, Université Laval, Québec, 1983.
- [7] A. Haurie and P. L'Écuyer, "Approximations and bounds in discrete event dynamic programming", les cahiers du GERAD, no. G-83-25, École des H.E.C. Montréal, 1983; to appear in *IEEE trans. on Automatic Control*.
- [8] M.L. Weitzman, "The ratchet principle and performance incentives", *Bell J. of Economics* 11 (1), 302-308 (1980).