Probability in the Engineering and Informational Sciences, **17**, 2003, 119–135. Printed in the U.S.A.

OPTIMAL CONTROL OF A SIMPLE IMMIGRATION PROCESS THROUGH THE INTRODUCTION OF A PREDATOR

E. G. KYRIAKIDIS

Department of Statistics and Actuarial Science
University of the Aegean
Karlovassi, 83200 Samos, Greece
E-mail: kyriak@aegean.gr

This article is concerned with the problem of controlling a simple immigration process, which represents a pest population, by the introduction of a predator. It is assumed that the cost rate caused by the pests is an increasing function of their population size and that the cost rate of the controlling action is constant. The existence of a control-limit policy that minimizes the expected long-run average cost per unit time is established. The proof is based on the variation of a fictitious parameter over the entire real line.

1. INTRODUCTION

The most widely used criteria for the optimal control of a stochastic process are the minimization of the expected total discounted cost and the minimization of the expected long-run average cost per unit time. Surveys of real applications of the relevant optimization models, in which the results of the studies have had some influence on actual decisions, are given in White [23,24]. In many problems, it is intuitively reasonable that the optimal policy initiates the controlling action if and only if the state of the process (e.g., number of customers in a queue, degree of deterioration of a machine, size of a biological population) exceeds a critical number. Such a policy is known as control-limit policy. The proof of the optimality of a control-limit policy is often an interesting problem and depends on the particular model. It is usually

© 2003 Cambridge University Press 0269-9648/03 $12.50

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

achieved by verifying that solutions to certain optimality equations possess specific properties such as monotonicity or convexity (see, e.g., Blackburn [4], Abakuks [1], Kawai [10], and Kyriakidis [11]). A standard method that sometimes leads to the proof of the optimality of a control-limit policy is the method of successive approximations. This method first establishes the optimality of a control-limit policy for the corresponding finite-horizon discounted cost problem. Then, this structural property is established for the infinite-horizon discounted cost problem, and, finally, to the average-cost problem. We refer to work by Derman [5], Douer and Yechiali [6], Van der Duyn Schouten and Vanneste [22], Benyamini and Yechiali [3], and Kyriakidis [13], in which this method is used. Federgruen and So [7] introduced a novel approach to establish the optimality of a control-limit policy by varying one of the problem parameters over the entire real line. First, it is shown that an optimal control-limit policy exists when the parameter takes sufficiently small values. This assertion is then extended inductively from interval to interval of the parameter values.

In this article, we will present a model for the optimal control of a damaging population through the introduction of a predator. The optimization criterion is the minimization of the average cost. The Federgruen–So technique is applied to establish the optimality of a control-limit policy. A difficulty that arises from the fact that the state space is infinite and the one-step expected costs are unbounded can be overcome by applying Sennott's [16] results. The method of successive approximations does not seem to be applicable.

In Section 2, the description of the model is given.

2. THE MODEL

Consider a pest population whose growth in a habitat is represented by a simple immigration process with immigration rate $\nu > 0$. It is assumed that the damage done by the pests is represented by a cost c_i, $i \geq 0$, for each unit of time during which the population size is i. It is natural to assume that the sequence $\{c_i\}$ is nondecreasing and $c_0 = 0$. Furthermore, we suppose that $c_i \to \infty$ as $i \to \infty$ and that $c_i \leq Ai^m$, $i \geq 1$, for some positive integer m and for some positive number A. The pest population may be controlled by some action that introduces a predator in the habitat after some random time that is exponentially distributed. The presence of the predator immediately stops the immigrations of the pests. As soon as the predator is introduced in the habitat, it captures the pests one at a time with constant rate $\sigma > 0$ until their population size is reduced to zero, and then it emigrates with rate $\theta > 0$. The unit of time has been chosen in such a way that the rate at which the predator is introduced in the habitat is equal to 1. Thus, when the controlling action is taken, the length of time until the introduction of the predator is exponentially distributed with unit mean. The controlling action gives rise to costs due to labor, materials, risk, and so forth. Let the cost of taking controlling action be K per unit time, where $K > 0$.

To illustrate the above assumptions, we could assume that the pest population consists of damaging insects, which immigrate to a habitat, and the predator is a lizard that may be set in the habitat after an exponentially distributed period of

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

time. Another realistic situation that reflects some elements of the model is the one in which the pests are seagulls that may be dangerous when an airplane takes off or lands at an airport near the sea and the predator is a hawk trained to chase them away.

Let i and i' be the states of the process at which the population size of the pests is i, $i \geq 0$, and the predator is absent from their habitat or present, respectively. A stationary policy f is defined by a sequence $\{f_i\}$, $i \geq 0$, where f_i is the action taken when the process is at state i. It is assumed that $f_i = 1$ when the controlling action is being taken and $f_i = 0$ when the controlling action is not being taken. If the stationary policy $f \equiv \{f_i\}$, $i \geq 0$, is employed, our assumptions imply that we have a continuous-time Markov chain model for the population growth of the pests with state space $S = \{0,0',1,1',\ldots\}$ and the following transition rates:

transition	rate
$i \to i+1$	ν $(i \geq 0)$
$i \to i'$	f_i $(i \geq 0)$
$i' \to (i-1)'$	σ $(i \geq 1)$
$0' \to 0$	θ

The expected long-run average cost per unit time of a policy π is defined as the limit as $t \to \infty$ of the expected cost incurred in the time interval $[0,t]$ divided by t, given that the policy π is employed. Our objective is to find a policy that minimizes the expected long-run average cost per unit time among all policies. The decision epochs include the epochs at which an immigration of a pest occurs and the epochs at which the predator emigrates. We place no restrictions on the class of allowable policies. Thus, the action chosen by a policy, when the process is at state $i \geq 0$, may depend on the history of the process or it may be randomized in the sense that it chooses action $a \in \{0,1\}$ with some probability that depends on a. An intuitively appealing class of policies is the class of control-limit policies $\{P_n : n = 0,1\ldots\}$, where P_n is the stationary policy under which the controlling action is taken if and only if the population size of the pests is equal to or exceeds n. It seems reasonable that the optimal policy will be of the control-limit type.

Our model can also have the following alternative interpretation. Suppose that customers arrive to a single-server queuing system according to a Poisson process with rate $\nu > 0$. The holding cost is equal to c_i when the number of customers is i. The status of the server can be either off or on. The controller may turn the server on and the preparations of this action last some time that is exponentially distributed with unit mean. The cost rate of the action equals $K > 0$. As soon as the server is on, the arrivals of the customers are halted. The service time for each customer is exponentially distributed with mean σ^{-1}. When the number of customers drops down to zero, the controller starts with preparations to switch the status from on to off. There are no switching costs, but the lead time of these preparations is exponentially distributed with mean θ^{-1}. The objective is to find the average-cost optimal policy, and the decision epochs include the epochs at which a customer arrives and the epochs at which the server becomes inactive. In this way, our model can be inter-

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

preted as a queuing model. We refer to Tijms' book [21, Chap. 3] and Sennott's book [17, Chap. 10] for various problems concerned with the optimal control of queuing systems.

In an earlier article (see Kyriakidis [11]), the above problem was studied in the special case in which $c_i = i$, $i \geq 0$. In that work, first, the optimal policy within the class of control-limit policies was found. The optimality of this policy within the wider class of all stationary policies was established by proving that it satisfies the optimality equation and certain conditions given by Bather [2].

It seems difficult to repeat this approach in the present model because the cost structure is more general and the average cost under a control-limit policy has a form that is too complicated. However, the existence of an optimal control-limit policy can be shown by applying a technique that was introduced by Federgruen and So [7–9] in some queuing and maintenance models. The same technique was developed in various replacement models (see So [18]), in stochastic models of a bottleneck facility (see So and Tang [19,20]), and in a different pest control model (see Kyriakidis [12]). In that work, it was assumed that the habitat to which the pests immigrate has a limited capacity, the immigration rates depend on their population size, and the controlling action introduces total catastrophes instead of a predator.

The Federgruen–So technique is based on a parametric analysis. In the present model, a fictitious parameter will be introduced and the Federgruen–So technique will be applied by varying this parameter. An important advantage of the Federgruen–So technique is that it enables us to show that the average cost of the control-limit policy P_n, $n \geq 0$, is unimodal as a function of n. This result permits us to obtain the optimal policy very quickly by implementing the standard bisection procedure or a tailor-made policy iteration algorithm operating on the class of control-limit policies.

The rest of the article is organized as follows. In Section 3, a fictitious parameter is introduced and the Federgruen–So technique is used to establish the optimality of a control-limit policy by varying this parameter over the entire real line. In Section 4, a discussion is presented about possible modifications of the model. In particular, we consider the case in which the sequence $\{c_i\}$ is bounded above, the case in which the predator may emigrate before capturing all the pests, and the case in which the pest population is controlled through total catastrophes.

3. THE OPTIMALITY OF CONTROL-LIMIT POLICIES

Let r be a real number (possibly negative) that represents a fictitious cost incurred each unit of time the process is occupying the state $0'$. In Theorem 6, it will be shown that a control-limit policy is optimal for any fixed value of r—in particular, for $r = 0$.

The present model is an infinite-state Markov decision process in continuous time. The existence of an optimal stationary policy cannot be proved by applying directly the classical results of Ross (see [15, Chap. 7]), because the one-step expected costs are unbounded. However, the transition rates and the cost structure of the present model are such that the general results of Sennott [16] can be easily applied. Sennott's Assumptions 1–5 are satisfied (see the details in the Appendix)

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

and, hence, according to her Theorem 2, it is deduced that an average-cost optimal stationary policy exists. The average cost g of the optimal policy does not depend on the initial state.

Let f be a stationary policy such that for every initial state, the state $0'$ can be reached with probability 1 after finitely many transitions. The average cost of the policy f is denoted by $g(f)$. Let also $T_{i0'}(f)$ $[T_{i'0'}(f)]$ and $C_{i0'}(f)$ $[C_{i'0'}(f)]$ be the expected times and costs, respectively, until the process reaches the state $0'$, given that the policy f is used and the initial state is i (i'). The quantity

$$h_i(f) = C_{i0'}(f) - g(f)T_{i0'}(f), \qquad i \geq 0,$$

is called the *relative value* (see Tijms [21, p. 188]) associated with the policy f, which corresponds to the state i. Similarly, the quantity

$$w_i(f) = C_{i'0'}(f) - g(f)T_{i'0'}(f), \qquad i \geq 0,$$

is the relative value associated with the policy f, which corresponds to the state i'. In the case in which $f = P_n$, instead of $g(f)$, $T_{i0'}(f)$, $T_{i'0'}(f)$, $C_{i0'}(f)$, $C_{i'0'}(f)$, $h_i(f)$, and $w_i(f)$ we use the notation g_n, $T_{i0'}^{(n)}$, $T_{i'0'}^{(n)}$, $C_{i0'}^{(n)}$, $C_{i'0'}^{(n)}$, $h_i^{(n)}$, and $w_i^{(n)}$, respectively.

We prove below that the optimal policy takes the controlling action that introduces the predator whenever the population size of the pests is sufficiently large.

LEMMA 1: *Let* $i^* = min\{i \geq 0 : c_i > g\}$. *The optimal policy prescribes action 1 for each* $i \geq i^*$.

PROOF: Assume that the optimal policy prescribes action 0 for some state $\tilde{i} \geq i^*$. Then, there exists an integer $i \geq i^*$ such that action 0 is prescribed in state i and action 1 is prescribed in state $i + 1$. (Otherwise, the predator would never be introduced in the habitat whenever the number of pests exceeds i^* so that the policy would have infinite cost.) Let h_j and w_j, $j \geq 0$, be the relative values associated with the optimal policy that correspond to the states j and j', respectively. Then,

$$\begin{aligned} h_i &= \frac{c_i - g}{\nu} + h_{i+1} \\ &= \frac{c_i - g}{\nu} + \frac{c_{i+1} + K - g + \nu h_{i+2} + w_{i+1}}{\nu + 1} \\ &\leq \frac{c_i + K - g + \nu h_{i+1} + w_i}{\nu + 1}, \end{aligned}$$

where the two equalities are based on relation (1) in Nobel and Tijms [14] and the inequality is due to the optimality of the policy. Hence,

$$\frac{c_i - g}{\nu} + \frac{(c_{i+1} - c_i) + \nu(h_{i+2} - h_{i+1}) + (w_{i+1} - w_i)}{\nu + 1} \leq 0. \qquad \textbf{(1)}$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

Note that

$$w_{i+1} - w_i = \left(\sum_{j=0}^{i+1} \frac{c_j}{\sigma} - g\,\frac{i+1}{\sigma} \right) - \left(\sum_{j=0}^{i} \frac{c_j}{\sigma} - g\,\frac{i}{\sigma} \right)$$
$$= \frac{c_{i+1} - g}{\sigma} > 0.$$

Let $T_{j0'}$ and $C_{j0'}$, $j \geq 0$, be the expected time and cost, respectively, until the process reaches the state $0'$, given that the optimal policy is employed and the initial state is j. Then,

$$h_{i+2} - h_{i+1} = C_{i+2,0'} - C_{i+1,0'} - g(T_{i+2,0'} - T_{i+1,0'})$$
$$\geq c_{i+1}(T_{i+2,0'} - T_{i+1,0'}) - g(T_{i+2,0'} - T_{i+1,0'})$$
$$= (c_{i+1} - g)(T_{i+2,0'} - T_{i+1,0'}) > 0.$$

The first of the above inequalities is explained by the monotonicity of $\{c_i\}$ and the fact that if the initial state is either $i+1$ or $i+2$, the process before entering the state $0'$ must pass through the state $(i+1)'$. Thus, the expression on the left-hand side of (1) is positive. This leads to a contradiction. ■

The lemma below gives some properties of a stationary policy that chooses action 1 whenever the number of pests is greater than or equal to a specific number. These properties will be used in the proof of Proposition 3.

LEMMA 2: *Suppose that the process is controlled by a stationary policy f, such that* $f(i) = 1$, $i \geq \hat{i}$. *Then, we have the following:*

(i) The quantities $C_{i0'}(f)$ *and* $C_{i'0'}(f)$, $i \geq \hat{i}$, *are bounded above by a polynomial in i.*

(ii) The equilibrium probabilities p_i, $i \geq \hat{i}$, *and* $p_{i'}$, $i \geq \hat{i}+1$, *of the corresponding embedded Markov chain have the following form:*

$$p_i = s_1 \left(\frac{\nu}{\nu+1} \right)^i, \qquad i \geq \hat{i},$$
$$p_{i'} = s_2 \left(\frac{\nu}{\nu+1} \right)^i, \qquad i \geq \hat{i}+1,$$

where s_1 *and* s_2 *are positive constants.*

PROOF: For $i \geq \hat{i}$, conditioning on the time until the introduction of the predator, we obtain

$$C_{i0'} = \int_0^\infty \left[\int_0^t E\{c_{X(s)} | X(0) = i\}\, ds \right] e^{-t}\, dt + \int_0^\infty \left[\sum_{j=i}^{\infty} p_{ij}(t) \sum_{k=1}^{j} \frac{c_k}{\sigma} \right] e^{-t}\, dt + K,$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

where $X(s)$ is the population size at time s of the (uncontrolled) simple immigration process and $p_{ij}(t)$ is the transition probability for this process. Using the condition $c_i \le Ai^m$ and after some calculations, we obtain the following result:

$$C_{i0'} \le \frac{A}{\nu+1}\sum_{\ell=0}^{m}\binom{m}{\ell} i^{\ell}\left[\sum_{n=0}^{\infty}\left(\frac{\nu}{\nu+1}\right)^n n^{m-l}\right] + \frac{A}{\sigma(\nu+1)}\sum_{\ell=0}^{m+1}\binom{m+1}{\ell} i^{\ell}\left[\sum_{n=0}^{\infty}\left(\frac{\nu}{\nu+1}\right)^n n^{m+1-\ell}\right] + K.$$

Note also that

$$C_{i'0'} = \sum_{j=1}^{i}\frac{c_j}{\sigma} \le \frac{1}{\sigma}\sum_{j=1}^{i} Aj^m \le \frac{Ai^{m+1}}{\sigma}.$$

Hence, part (i) of the lemma has been proved. To prove part (ii), note that the process under the policy f is stable and its stationary probabilities π_i, $i \ge \hat{i}-1$, and $\pi_{i'}$, $i \ge 0$, satisfy the equations

$$(\nu+1)\pi_i = \nu\pi_{i-1}, \qquad i \ge \hat{i},$$
$$\nu\pi_i = \sigma\pi_{(i+1)'}, \qquad i \ge 0.$$

From the above equations, it follows that

$$\pi_i = \left(\frac{\nu}{\nu+1}\right)^{i+1-\hat{i}}\pi_{\hat{i}-1}, \qquad i \ge \hat{i},$$
$$\pi_{i'} = \frac{\nu}{\sigma}\left(\frac{\nu}{\nu+1}\right)^{i-\hat{i}}\pi_{\hat{i}-1}, \qquad i \ge \hat{i}+1.$$

From the above expressions, the desired form of the equilibrium probabilities of the corresponding embedded Markov chain can be obtained (see relation (2.6.1) in Tijms [21]). ■

PROPOSITION 3: *The policy P_n, $n \ge 0$, is optimal if and only if*

$$\frac{c_i - g_n}{\nu} + h_{i+1}^{(n)} \le \frac{c_i + K - g_n + \nu h_{i+1}^{(n)} + w_i^{(n)}}{\nu+1}, \qquad 0 \le i \le n-1, \tag{2}$$

and

$$\frac{c_i + K - g_n + \nu h_{i+1}^{(n)} + w_i^{(n)}}{\nu+1} \le \frac{c_i - g_n}{\nu} + h_{i+1}^{(n)}, \qquad i \ge n. \tag{3}$$

PROOF: Suppose that the policy P_n is optimal. According to relation (1) in Nobel and Tijms [14], the numbers $h_i^{(n)}$ and $w_i^{(n)}$, $i \ge 0$, and g_n satisfy the following system of linear equations:

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

$$h_i^{(n)} = \frac{c_i - g_n}{\nu} + h_{i+1}^{(n)}, \qquad 0 \le i \le n-1,$$

$$h_i^{(n)} = \frac{c_i + K - g_n + \nu h_{i+1}^{(n)} + w_i^{(n)}}{\nu + 1}, \qquad i \ge n,$$

$$w_i^{(n)} = \frac{c_i - g_n}{\sigma} + w_{i-1}^{(n)}, \qquad i \ge 1,$$

$$w_0^{(n)} = \frac{r - g_n}{\theta} + h_0^{(n)}.$$

Assume that for some $\hat{i}$, $0 \le \hat{i} \le n-1$, inequality (2) does not hold. Then, in view of relation (2) in Nobel and Tijms [14], the stationary policy f with

$$f(i) = 0, \qquad 0 \le i \le n-1, i \ne \hat{i},$$
$$f(\hat{i}) = 1,$$
$$f(i) = 1, \qquad i \ge n,$$

has a smaller average cost than the average cost of P_n. This is a contradiction. Hence, inequality (2) holds. Similarly, it can be shown that inequality (3) is valid.

Suppose that the policy P_n satisfies inequalities (2) and (3). Hence, the numbers g_n, $h_i^{(n)}$, and $w_i^{(n)}$, $i \ge 0$, satisfy the following equations:

$$h_i^{(n)} = \min\left\{\frac{c_i - g_n}{\nu} + h_{i+1}^{(n)}, \frac{c_i + K - g_n + \nu h_{i+1}^{(n)} + w_i^{(n)}}{\nu + 1}\right\}, \qquad i \ge 0, \qquad \textbf{(4)}$$

$$w_i^{(n)} = \frac{c_i - g_n}{\sigma} + w_{i-1}^{(n)}, \qquad i \ge 1,$$

$$w_0^{(n)} = \frac{r - g_n}{\theta} + h_0^{(n)}.$$

Let f be a stationary policy such that $f(i) = 1$, $i \ge i^*$. In view of Lemma 1, to prove that the policy P_n is overall optimal, it suffices to show that, for every initial state i, its average cost is smaller than or equal to the average cost of the policy f.

Let X_N and a_N be the state of the process and the action chosen, respectively, after the Nth ($N \ge 1$) transition. We assume that X_0 and a_0 are the initial state and initial action, respectively. Also, let $T(i,a)$ and $C(i,a)$ be the expected time and cost, respectively, until a transition occurs when action a is chosen in state i. From the proof of Theorem 7.6 in Ross [15], it follows that

$$g_n \le \frac{E_f[u(X_N) - u(X_0)] + E_f \sum_{i=1}^{N} C(X_{i-1}, a_{i-1})}{E_f \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})}, \qquad \textbf{(5)}$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

where $u(X_N)$ is equal to $h_{X_N}^{(n)}$ if $X_N \in \{0,1,\ldots\}$ or to $w_{X_N}^{(n)}$ if $X_N \in \{0',1',\ldots\}$, and E_f represents the conditional expectation given that the policy f is employed. Relation (5) holds with equality for $f = P_n$. Using the definition $u(X_N)$, relation (5) can be written as

$$g_n \leq \frac{E_f C_{X_N 0'}^{(n)} - g_n E_f T_{X_N 0'}^{(n)} - u(X_0) + E_f \sum_{i=1}^{N} C(X_{i-1}, a_{i-1})}{E_f \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})},$$

or, equivalently,

$$g_n \leq \frac{\sum_{j \in S} p_{X_0 j}^{(N)} C_{j0'}^{(n)} - g_n \sum_{j \in S} p_{X_0 j}^{(N)} T_{j0'}^{(n)} - u(X_0) + E_f \sum_{i=1}^{N} C(X_{i-1}, a_{i-1})}{E_f \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})},$$

where $p_{X_0 j}^{(N)}$ is the probability that the state of the process under the policy f after N transitions will be j. Letting $N \to \infty$ and using the bounded convergence theorem (see Tijms [21, p. 351]), we can interchange the order of limit and summation in the numerator of the above expression. Hence,

$$g_n \leq \frac{\sum_{j \in S} p_j C_{j0'}^{(n)} - g_n \sum_{j \in S} p_j T_{j0'}^{(n)} - u(X_0)}{\lim_{N \to \infty} E_f \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})} + \lim_{N \to \infty} \frac{E_f \sum_{i=1}^{N} C(X_{i-1}, a_{i-1})}{E_f \sum_{i=1}^{N} T(X_{i-1}, a_{i-1})},$$

where p_j, $j \in S$, are the equilibrium probabilities of the corresponding embedded Markov chain of the process under the policy f. Using Lemma 2, it is deduced that the numerator of the first term of the expression on the right-hand side in the above relation is finite. Since the denominator, as $N \to \infty$, tends to infinity, we deduce that this term is zero. The second term is the expected long-run average cost per unit time of the policy f. Hence, $g_n \leq g_f(X_0)$, and, thus, the policy P_n is optimal. ■

Let g_n denote the expected long-run average cost per unit time under the policy P_n, $n \geq 0$. The process under the policy P_n is a regenerative process, where the successive entries into the state $0'$ can be taken as regenerative epochs between successive cycles. From a well-known regenerative argument (see Ross [15, Prop. 5.9]), it follows that g_n is equal to the expected cost of a cycle divided by the expected time of a cycle. Hence,

$$g_n = \left(\frac{r}{\theta} + \sum_{i=0}^{n-1} \frac{c_i}{\nu} + C_{n0'}^{(n)} \right) \left(\frac{1}{\theta} + \frac{n}{\nu} + T_{n0'}^{(n)} \right)^{-1}. \tag{6}$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

For $n = 0, 1, \ldots$, let

$$A_i^{(n)} = C_{i+1,0'}^{(n)} - \sigma^{-1} \sum_{j=1}^{i} c_j, \qquad i \geq n.$$

The results of Lemmas 4 and 5 will be used in the proof of Theorem 6.

LEMMA 4: *For all $n \geq 0$, the sequence $\{A_i^{(n)}\}$, $i \geq n$, is nondecreasing in i.*

PROOF: The sequence $\{A_i^{(n)}\}$, $i \geq n$, is nondecreasing in i if and only if

$$C_{i+2,0'}^{(n)} - C_{i+1,0'}^{(n)} \geq \frac{c_{i+1}}{\sigma}, \qquad i \geq n.$$

Conditioning on the first transition from the state $i+1$, we can obtain an expression of $C_{i+1,0'}^{(n)}$ in terms of $C_{i+2,0'}^{(n)}$. Thus, the above relation takes the following form:

$$C_{i+1,0'}^{(n)} \geq c_i + K + \sum_{j=1}^{i} \frac{c_j}{\sigma} + \frac{(\nu+1)c_i}{\sigma}, \qquad i \geq n+1. \tag{7}$$

Conditioning on the time until the introduction of the predator, we have

$$C_{i+1,0'}^{(n)} = \int_0^\infty \left[\int_0^t E\{c_{X(s)} | X(0) = i+1\} ds \right] e^{-t}\, dt + \int_0^\infty \left[\sum_{j=i+1}^{\infty} p_{i+1,j}(t) \sum_{k=1}^{j} \frac{c_k}{\sigma} \right] e^{-t}\, dt + K,$$

where $X(s)$ is the population size at time s of the (uncontrolled) simple immigration process and $p_{i+1,j}(t)$ is the transition probability for this process. The above expression can be written as

$$C_{i+1,0'}^{(n)} = \sum_{j=i+1}^{\infty} p_{i+1,j}^* c_j + \sum_{j=i+1}^{\infty} p_{i+1,j}^* \left(\sum_{k=1}^{j} \frac{c_k}{\sigma} \right) + K,$$

where $p_{i+1,j}^*$ is the Laplace transform of $p_{i+1,j}(t)$ at the point 1. From the monotonicity of $\{c_i\}$, it follows that

$$\sum_{j=i+1}^{\infty} p_{i+1,j}^* c_j \geq \sum_{j=i+1}^{\infty} p_{i+1,j}^* c_i = c_i.$$

Hence, (7) holds if the following inequality holds:

$$\sum_{j=i+1}^{\infty} p_{i+1,j}^* \left(\sum_{k=1}^{j} c_k \right) \geq (\nu+1)c_i + \sum_{j=1}^{i} c_j, \qquad i \geq n+1.$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

The above relation simplifies to

$$\sum_{j=i+1}^{\infty} p^*_{i+1,j}\left(\sum_{k=i+1}^{j} c_k\right) \geq (\nu+1)c_i, \qquad i \geq n. \tag{8}$$

Note that

$$\begin{aligned}\sum_{j=i+1}^{\infty} p^*_{i+1,j}\left(\sum_{k=i+1}^{j} c_k\right) &\geq \sum_{j=i+1}^{\infty} p^*_{i+1,j}(j-i)c_{i+1} \\ &= c_{i+1}\sum_{j=i+1}^{\infty}(j-i)\int_0^{\infty}\frac{(\nu t)^{j-i-1}}{(j-i-1)!}e^{-(\nu+1)t}\,dt \\ &= (\nu+1)c_{i+1}.\end{aligned}$$

Hence, (8) holds and the proof of the lemma is complete. ■

LEMMA 5: *Assume that the policy P_n, $n>0$, is optimal for some fixed value R of the parameter r. Then, it is impossible for the policy P_n to be optimal for all $r \geq R$.*

PROOF: Suppose that P_n is optimal for all $r \geq R$. Then, the numbers $h_i^{(n)}$ and $w_i^{(n)}$, $i \geq 0$, satisfy (3) at $i=n$ for all $r \geq R$. Hence,

$$\frac{c_n + K - g_n + \nu h_{n+1}^{(n)} + w_n^{(n)}}{\nu+1} \leq \frac{c_n - g_n}{\nu} + h_{n+1}^{(n)}.$$

Using the definition of the relative values $h_{n+1}^{(n)}$ and $w_n^{(n)}$, the above relation reduces to

$$[\nu(T_{n+1,0'}^{(n)} - T_{n'0'}^{(n)}) + 1]g_n \leq c_n + \nu(C_{n+1,0'}^{(n)} - C_{n'0'}^{(n)}) - \nu K.$$

From (6), it follows that $g_n \to \infty$ as $r \to \infty$. This contradicts the above inequality because its right-hand side does not depend on r and the coefficient of g_n is positive. ■

We are now ready to prove our main result—the optimality of control-limit policies.

THEOREM 6: *There exists an increasing sequence $\{R_n\}$, $n \geq 0$, with $R_0 = -\infty$, $R_1 > -\infty$, $R_n < \infty$, and $\lim_{n\to\infty} R_n = +\infty$ such that the policy P_n, $n \geq 0$, is optimal for all $r \in [R_n, R_{n+1}]$, where*

$$R_{n+1} = \sup\{w : w \geq R_n \text{ and the policy } P_n \text{ is optimal for all } r \in [R_n, w]\}.$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

PROOF: The proof is, by induction, on n. We fix a value r_0 of r. From Lemma 1, it follows that the optimal policy chooses only the action 1 for all $i \geq i^*(r_0)$, where

$$i^*(r_0) = \min\{i \geq 1 : c_i > g(r_0)\}$$

and $g(r_0)$ is the minimum average cost. Let $r \leq r_0$. The critical number $i^*(r)$ is nondecreasing in r, $r \leq r_0$, since $g(r)$ is nondecreasing in r. Hence, in view of Proposition 3, the policy P_0 is optimal for $r \leq R$ with $R \leq r_0$ if the numbers $h_i^{(0)}$ and $w_i^{(n)}$, $i \geq 0$, and g_0 satisfy the inequality

$$\frac{c_i + K - g_0 + \nu h_{i+1}^{(0)} + w_i^{(0)}}{\nu + 1} \leq \frac{c_i - g_0}{\nu} + h_{i+1}^{(0)}, \qquad 0 \leq i \leq i^*(r_0).$$

Using the definition of the relative values $h_{i+1}^{(0)}$ and $w_i^{(0)}$, $i \geq 0$, the above inequality reduces to

$$g_0[\nu(T_{i+1,0'}^{(0)} - T_{i'0'}^{(0)}) + 1] \leq c_i + \nu(C_{i+1,0'}^{(0)} - C_{i'0'}^{(0)}) - \nu K, \qquad 0 \leq i \leq i^*(r_0). \quad \textbf{(9)}$$

From (6), we have that $g_0 \to -\infty$ as $r \to -\infty$. Thus, there exist a number $R > -\infty$ with $R \leq r_0$ such that inequality (9) holds simultaneously for all $r \leq R$. From Lemma 5, it follows that $R_1 < +\infty$, where $R_1 = \sup\{w : w \geq R$ and the policy P_0 is optimal for all $r \leq w\}$.

Suppose that there exists a sequence $R_0 < R_1 < \cdots < R_n$ such that the policy P_s, $1 \leq s \leq n$, is optimal for all $r \in [R_s, R_{s+1}]$ with $R_{n+1} = \sup\{w : w \geq R_n$ and the policy P_n is optimal for all $r \in [R_n, w]\} < +\infty$. We will show that the policy P_{n+1} is optimal for $r = R_{n+1}$. Consider some $\epsilon > 0$. If $r = R_{n+1} + \epsilon$, the policy P_n is not optimal. From Proposition 3, it follows that one of the following two cases occurs.

Case 1: For some i with $0 \leq i \leq n-1$,

$$\frac{c_i + K - g_n + \nu h_{i+1}^{(n)} + w_i^{(n)}}{\nu + 1} < \frac{c_i - g_n}{\nu} + h_{i+1}^{(n)}.$$

The above inequality is equivalent to $\psi_i(R_{n+1} + \varepsilon) > 0$, with

$$\begin{aligned}\psi_i(r) &= c_i - g_n + \nu h_{i+1}^{(n)} - \nu w_i^{(n)} - \nu K \\ &= c_i + \nu\left(C_{i+1,0'}^{(n)} - \sum_{j=1}^{i} \frac{c_j}{\sigma}\right) - g_n\left[\nu\left(T_{i+1,0'}^{(n)} - \frac{i}{\sigma}\right) + 1\right] - \nu K.\end{aligned}$$

Since g_n, as given in (6), is increasing in r, we deduce that $\psi_i(r)$ is decreasing in r. Thus,

$$0 < \psi_i(R_{n+1} + \epsilon) < \psi_i(R_{n+1}) \leq 0,$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

where the last inequality follows from the optimality of P_n for $r = R_{n+1}$. Clearly, this is a contradiction and Case 2 must arise.

Case 2: For some i with $i \geq n$,

$$\frac{c_i - g_n}{\nu} + h_{i+1}^{(n)} < \frac{c_i + K - g_n + \nu h_{i+1}^{(n)} + w_i^{(n)}}{\nu + 1}.$$

The above inequality is equivalent to $\psi_i(R_{n+1} + \varepsilon) < 0$ with

$$\psi_i(r) = c_i + \nu\left(C_{i+1,0'}^{(n)} - \sum_{j=1}^{i} \frac{c_j}{\sigma}\right) - g_n\left[\nu\left(T_{i+1,0'}^{(n)} - \frac{i}{\sigma}\right) + 1\right] - \nu K$$

$$= c_i + \nu A_i^{(n)} - g_n\left[\nu\left(1 + \frac{\nu + 1}{\sigma}\right) + 1\right] - \nu K,$$

where the last inequality follows from the expression $T_{i+1,0'}^{(n)} = 1 + (\nu + i + 1)\sigma^{-1}$, which was derived in Kyriakidis [11]. From Lemma 4 and the monotonicity of $\{c_i\}$, it follows that the sequence $\psi_i(r)$, $i \geq n$, is nondecreasing in i. Thus,

$$\psi_n(R_{n+1} + \varepsilon) \leq \psi_i(R_{n+1} + \varepsilon) < 0.$$

Consider a sequence $\{\varepsilon_k\} \downarrow 0$. In view of the above inequality, we have that for all k, $\psi_n(R_{n+1} + \varepsilon_k) < 0$. From the continuity of $\psi_n(r)$ in r, it follows that $\psi_n(R_{n+1}) \leq 0$. However, $\psi_n(R_{n+1}) \geq 0$ because the policy P_n is optimal for $r = R_{n+1}$. Thus, $\psi_n(R_{n+1}) = 0$. The last equality means that for $r = R_{n+1}$, the right-hand side of (4) attains the same value at the state $i = n$ for both states 0 and 1. Therefore, the action prescribed, for each state $i \geq 0$, by the policy P_{n+1} minimizes the right-hand side of (4). Thus, the policy P_{n+1} is optimal for $r = R_{n+1}$. Define

$$R_{n+2} = \sup\{w : w \geq R_{n+1} \text{ and the policy } P_{n+1} \text{ is optimal for all } r \in [R_{n+1}, w]\}.$$

From Lemma 5, it follows that $R_{n+2} < +\infty$. It remains to be shown that $\lim_{n\to\infty} R_n = +\infty$. Assume to the contrary that $\lim_{n\to\infty} R_n = R^* < +\infty$. Choose $n^* > i^*(R^*)$. From Lemma 1, it follows that for $r = R_{n^*}$, the optimal policy chooses only action 1 in some state i with $n^* > i \geq i^*(R^*) \geq i^*(R_{n^*})$. This contradicts the optimality of the policy P_{n^*} for $r = R_{n^*}$. ■

LEMMA 7: $\{T_{n0'}^{(n)}\}$, $n \geq 0$, *is nondecreasing in* n.

PROOF: Conditioning on the first transition from state n, we have that

$$T_{n0'}^{(n)} = \frac{1}{\nu + 1} + \frac{\nu}{\nu + 1} T_{n+1,0'}^{(n+1)} + \frac{n}{(\nu + 1)\sigma}.$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

Hence,

$$T_{n+1,0'}^{(n+1)} - T_{n0'}^{(n)} = \frac{1}{\nu+1}\left(T_{n+1,0'}^{(n+1)} - 1 - \frac{n}{\sigma}\right). \tag{10}$$

Conditioning on the time until the introduction of the predator, we obtain

$$\begin{aligned} T_{n+1,0'}^{(n+1)} &= \int_0^\infty \left[t + \sum_{j=n+1}^{\infty} p_{n+1,j}(t)\,\frac{j}{\sigma}\right] e^{-t}\,dt \\ &\geq 1 + \int_0^\infty \left[\sum_{j=n+1}^{N} p_{n+1,j}(t)\,\frac{n}{\sigma}\right] e^{-t}\,dt \\ &= 1 + \frac{n}{\sigma}, \end{aligned} \tag{11}$$

where $p_{n+1,j}(t)$ is the probability that the state of the (uncontrolled) simple immigration process at time t will be j, given that the state at time 0 is $n+1$. Relations (10) and (11) give the result of the lemma. ■

From (6) and Lemma 7, we deduce that g_n can be written as

$$g_n = r\pi_n + \tilde{g}_n,$$

where $\tilde{g}_n$ is independent of r and π_n is decreasing in n. Using this result, the following proposition, which is useful for the computation of the optimal control-limit policy, can be proved in the same way as Lemma 4 in Kyriakidis [12].

PROPOSITION 8: *For any fixed r, the average cost g_n, $n \geq 0$, is unimodal as a function of n.*

In view of Proposition 8, the critical point n^* of the optimal control-limit policy P_{n^*} can be determined either by the standard bisection procedure, which determines the minimum within the set $\{g_n : n \geq 0\}$ or by a tailor-made policy iteration algorithm (see Tijms [21], Nobel and Tijms [14], and Kyriakidis [12]), which generates a sequence of strictly improved control-limit policies and converges to P_{n^*}.

4. MODIFICATIONS OF THE MODEL

Suppose that the sequence $\{c_i\}$ is bounded above. In this case, the average cost of the policy that never takes the controlling action that introduces the predator is equal to $\sup_i c_i$. This policy is optimal if and only if

$$\frac{c_i - \sup_i c_i}{\nu} + h_{i+1} \leq \frac{c_i + K - \sup_i c_i + \nu h_{i+1} + w_i}{\nu+1}, \qquad i \geq 0, \tag{12}$$

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

where the sequences $\{h_i\}$ and $\{w_i\}$, $i \geq 0$, satisfy the difference equations

$$h_i = \frac{c_i - \sup_i c_i}{\nu} + h_{i+1}, \qquad i \geq 0,$$

$$w_i = \frac{c_i - \sup_i c_i}{\sigma} + w_{i-1}, \qquad i \geq 1,$$

$$w_0 = -\frac{\sup_i c_i}{\theta} + h_0.$$

It can be checked that (12) holds if and only if

$$(\nu^{-1} + \sigma^{-1}) \sum_{j=1}^{\infty} (\sup_i c_i - c_j) + (\nu^{-1} + \theta^{-1}) \sup_i c_i \leq K. \tag{13}$$

Suppose that (13) does not hold. Then, it can be shown that the optimal policy is of control-limit type by applying the Federgruen–So technique as in the case in which $c_i \to \infty$, as $i \to \infty$.

Consider now the case in which the predator may emigrate with rate $\theta > 0$ from the habitat before capturing all the pests; that is, the transition $i' \to i (i \geq 0)$ occurs with rate θ. There is strong numerical evidence from various numerical examples that the optimal policy is again of control-limit type. However, it seems difficult to prove this conjecture. In the special case in which $c_i = i$, $i \geq 0$, it is possible to show that the policy P_0 is optimal among all stationary policies if $\sigma - \nu\theta > 0$ and $k \leq \nu\theta^{-1}(\nu + 1 + \sigma + \theta)(\sigma - \nu\theta)^{-1}$. The proof is based on Bather's [2] general results following the same procedure as in Kyriakidis [11].

Finally, we consider the case in which the controlling action, instead of a predator, introduces total catastrophes. When such a catastrophe occurs, the population size of the pests is instantaneously reduced to zero. In a previous article [13], we considered this model with the assumption that the sequence $\{c_i\}$ is bounded above. In that work, the optimality of a control-limit policy was first established for the corresponding finite-horizon discounted cost problem. This structural property was then transferred to the infinite-horizon discounted cost problem and, finally, to the average cost problem. However, the method that we followed in Section 3 can be used in this model in order to establish the optimality of control-limit policies, even in the case in which the sequence $\{c_i\}$ is unbounded.

References

1. Abakuks, A. (1979). An optimal hunting policy for a stochastic logistic model. *Journal of Applied Probability* 16: 319–331.
2. Bather, J. (1976). Optimal stationary policies for denumerable Markov chains in continuous time. *Advances in Applied Probability* 8: 144–158.

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

3. Benyamini, Z. & Yechiali, U. (1999). Optimality of control limit maintenance policies under non-stationary deterioration. *Probability in the Engineering and Informational Sciences* 13: 55–70.
4. Blackburn, J.D. (1972). Optimal control of a single-server queue with balking and reneging. *Management Science* 19: 297–313.
5. Derman, C. (1963). On optimal replacement rules when changes of state are Markovian. In R. Bellman (ed.), *Mathematical optimization techniques*. Berkeley: University of California Press, pp. 201–210.
6. Douer, N. & Yechiali, U. (1972). Optimal repair and replacement in Markovian systems. *Communications in Statistics—Stochastic Models* 10(1): 253–270.
7. Federgruen, A. & So, K.C. (1989). Optimal time to repair a broken server. *Advances in Applied Probability* 21: 709–731.
8. Federgruen, A. & So, K.C. (1990). Optimal maintenance policies for single-server queueing systems subject to breakdowns. *Operations Research* 38: 330–343.
9. Federgruen, A. & So, K.C. (1991). Optimality of threshold policies in single-server queueing systems with server vacations. *Advances in Applied Probability* 23: 388–405.
10. Kawai, H. (1983). An optimal ordering and replacement policy of a Markovian degradation system under complete observation—Part I. *Journal of Operational Research Society of Japan* 26: 279–290.
11. Kyriakidis, E.G. (1995). Optimal pest control through the introduction of a predator. *European Journal of Operational Research* 81: 357–363.
12. Kyriakidis, E.G. (1999). Optimal control of a truncated general immigration process through total catastrophes. *Journal of Applied Probability* 36: 461–472.
13. Kyriakidis, E.G. (1999). Characterization of the optimal policy for the control of a simple immigration process through total catastrophes. *Operations Research Letters* 24: 245–248.
14. Nobel, R.D. & Tijms, H.C. (1999). Optimal control of an $M^X/G/1$ queue with two service modes. *European Journal of Operational Research* 113: 610–619.
15. Ross, S.M. (1992). *Applied probability models with optimization applications*. New York: Dover.
16. Sennott, L.I. (1989). Average cost semi-Markov decision processes and the control of queueing systems. *Probability in the Engineering and Informational Sciences* 3: 247–272.
17. Sennott, L.I. (1999). *Stochastic dynamic programming and the control of queueing systems*. New York: Wiley.
18. So, K.C. (1992). Optimality of control-limit policies in replacement models. *Naval Research Logistics* 39: 685–697.
19. So, K.C. & Tang, C.S. (1995). Optimal operating policy for a bottleneck with random rework. *Management Science* 41: 620–636.
20. So, K.C. & Tang, C.S. (1995). Optimal batch sizing and repair strategies for operations with repairable jobs. *Management Science* 41: 894–908.
21. Tijms, H.C. (1994). *Stochastic models: An algorithmic approach*. New York: Wiley.
22. Van der Duyn Schouten, F.A. & Vanneste, S.G. (1995). Maintenance optimization of a production system with buffer capacity. *European Journal of Operational Research* 82: 323–338.
23. White, D.J. (1985). Real applications of Markov decision processes. *Interfaces* 15: 73–83.
24. White, D.J. (1988). Further real applications of Markov decision processes. *Interfaces* 18: 55–61.

APPENDIX

We shall show that an optimal stationary policy exists in our model using the general results by Sennott [16]. We first introduce some general notation and five assumptions for a semi-Markov decision process.

The notation used is as follows:

S	State space
$p_{ij}(a)$	Transition probability to state j if action a is taken in state i

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press

$T(i,a)$	One-step expected time if action a is taken in state i
$C(i,a)$	One-step expected cost if action a is taken in state i
$d(i,a)$	Cost per unit time if action a is taken in state i
$d(i)$	$\min_a\{d(i,a)\}$
$c_{ij}(f)$	Expected cost of a first passage from the state i to state j under a stationary policy f
$\alpha > 0$	Discount factor
$V_{\pi,\alpha}(i)$	Total expected α—discounted cost under the policy π with an initial state i
$V_\alpha(i)$	$\inf_\pi V_{\pi,\alpha}(i)$
$h_\alpha(i)$	$V_\alpha(i) - V_\alpha(0)$, where 0 is any distinguished state.

ASSUMPTION 1: *There exists $\delta > 0$ and $\varepsilon > 0$ such that, for every state and action, there is a probability of at least ε that the transition time will be greater than δ.*

ASSUMPTION 2: *There exists B such that $T(i,a) \le B$ for every i and a.*

ASSUMPTION 3: *$V_\alpha(i) < +\infty$ for all i and α.*

ASSUMPTION 4: *There exists $\alpha_0 > 0$ and nonnegative numbers M_i such that $h_\alpha(i) \le M_i$ for every i and $0 < \alpha < \alpha_0$. For every i, there exists an action $a(i)$ such that $\sum_j p_{ij}(a(i))M_j < \infty$.*

ASSUMPTION 5. *There exists $\alpha_0 > 0$ and a nonnegative number N such that $-N \le h_\alpha(i)$ for every i and $0 < \alpha < \alpha_0$.*

Sennott proved that if Assumptions 1–5 hold, then there exists a stationary policy that is average-cost optimal. Sennott also showed the following assertions:

(a) Assumptions 3 and 4 hold if Assumption 1 holds and there exists a stationary policy f that induces an irreducible ergodic embedded Markov chain on S and

$$\sum_{i\in S} p_i C(i, f(i)) < \infty, \tag{A1}$$

$$\sum_{i\in S} p_i T(i, f(i)) < \infty, \tag{A2}$$

where p_i, $i \in S$, are the steady-state probabilities of the embedded Markov chain.

(b) Assumption 5 holds if there exists a stationary policy f with average cost b, which satisfies the assumptions in (a) and the following:
   (i) There exists $\varepsilon > 0$ and a finite subset G of S such that $d(i) \ge b + \varepsilon$ for $i \in S - G$.
   (ii) For each $i \in G$, there exists a stationary policy f_i such that $c_{0i}(f_i) < \infty$.

In the present model, Assumption 1 holds with $\delta = 1$, $\varepsilon = \min\{\exp[-(\nu+1)], \exp(-\sigma), \exp(-\theta)\}$ and Assumption 2 holds with $B = \max\{\sigma^{-1}, \theta^{-1}, \nu^{-1}\}$. Assumptions 3 and 4 hold because Assertion (a) can be used with the stationary policy $f = P_0$. Note that, in view of Lemma 2(ii), (A1) and (A2) are true. Assumption 5 holds because Assertion (b) can be used with $f = P_0$, $b = g_0$, $G = \{0,0',1,1',\ldots,j,j'\}$ with $c_j \ge b + \varepsilon$, $\varepsilon = 1$, and $f_i = P_j (i \in G)$.

https://doi.org/10.1017/S0269964803171070 Published online by Cambridge University Press