No.605

A SEQUENTIAL ALLOCATION PROBLEM WITH SEARCH COST WHERE THE SHOOT-LOOK-SHOOT POLICY IS EMPLOYED

by

Masahiro Sato Seizo Ikuta

October 1994

A SEQUENTIAL ALLOCATION PROBLEM WITH SEARCH COST WHERE THE SHOOT-LOOK-SHOOT POLICY IS EMPLOYED

Masahiro Sato   Seizo Ikuta
University of Tsukuba

1994.10.4

Abstract Suppose a hunter starts hunting over t periods with i bullets. A distribution of the value of targets appearing and the hit probability of a bullet are known. For shooting, he takes a strategy of a shoot-look-shoot scheme. The objective in this article is to find an optimal decision policy which maximizes his total expected present discounted reward. In the case with no search cost, the optimal policy is monotone in the number of bullets remaining then, but not always monotone in the case with search cost.

1 Introduction

This paper presents a stochastic dynamic programming model of a sequential allocation problem with search cost. In the problem, we invest some units of resource, one by one, for a present opportunity whose value is a random sample from a known probability distribution. The one by one unit investment is often called a *shoot-look-shoot* policy, implying that, if unsuccessful after investing one unit, it is decided whether to invest one more unit or not. This problem can be applied to the following examples.

• **Hunting Problem** Suppose a hunter sees a target of value w. Then, he must decide whether to shoot it or not. If he decides not to shoot, then he must search for another target. On the other hand, suppose he decides to shoot. Then, if the bullet hits, he can get the target, and if it does not hit, the target may escape. When it does not escape, he must decide whether to shoot another bullet or not.

• **Advertising Problem** Suppose a salesman visits different stores with i samples to advertise his new products. Customers who come to a store will try the samples. Seeing the customers' responses, the manager of the store decides whether to sell it in his store or not. If the manager decides to sell, then it may be regarded as a success and the salesman can obtain a profit w and begin searching for another store, or else he must search for another store with no profit from the previous store. It may be possible that the manager still has not decided even after some customers have tried the samples. Then, the salesman has to decide whether to continue advertising in the store or to quit and search for another store.

By the way, if an opportunity can be successfully achieved by investing only one unit of resource, the problem can be reduced to an optimal stopping problem with no recall where i opportunities can be obtained.

Sequential allocation problems have been discussed by many authors so far. Derman, Lieberman and Ross[1], and Prastacos[7] considered them as investment problems in which the expected reward depended on the amount of resources that were invested. Mastran and Thomas[5] treated them as target attacking problems where the decision policy of a shoot-look-shoot scheme was discussed. They showed the computation method for the optimal decision rule, but did not mathematically verify its structure. Kisi[4] considered a similar model with a shoot-look-shoot policy where the relation between approximate solutions and exact ones were mainly discussed. Sakaguchi[8] investigated the continuous-time version model of [5] in which a case of a shoot-look-shoot scheme was also discussed. He derived the conclusion that the critical value, at which shooting and not shooting become indifferent in the optimal decision, was nonincreasing in the number of remaining torpedoes. Namekata, Tabata and Nishida[6] also dealt with a model similar to [5] where there exist two kinds of targets in a sense that the necessary number of torpedoes to get the targets is different. Revealing the structure of the optimal decision policy, they did not discuss about the shoot-look-shoot policy. Now, it should be noted that in all of the models above, a search cost was not introduced. However, among the variations of the above models there exist ones in which a search cost must be assumed necessary.

The objectives of the present paper are to pose a general model in which a search cost is an essential factor and to examine properties of the optimal decision policy.

One of the most distinctive results obtained in the present paper is that *the critical value does not always become nonincreasing in the number of remaining units of resource.*

In Section 2 that follows, we define our model and in Section 3 its fundamental equations are derived. In Section 4 the structure of the optimal decision policy is investigated. In Sections 5, 6 and 7, we state cases with no search cost, with search cost and with a sufficiently large search cost, respectively. In Section 8 the case of an infinite amount of available resources is examined and in Section 9 some numerical examples are given. In Section 10 the conclusions obtained are summarized. Finally, some limitations of the present model are stated in Section 11.

2 Model

Throughout this paper, we shall explain our problem with the following hunting problem. Suppose a hunter starts hunting over t periods with i bullets. In order to go shooting at a certain time, he must pay a search cost c at the previous point in time. In the woods, he can find a target with a probability $\theta \in (0,1]$, assuming that more than one target can not be found at any point in time. The value of the target that is found is a random variable having a known probability distribution function $F_1(w)$, continuous or discrete, where $F_1(w) = 0$ for $w < a < 1$, $0 < F_1(w) < 1$ for $0 < a \leq w < 1$, and $F_1(w) = 1$ for $1 \leq w$. The values of the targets that have been found at the successive points in time are assumed to be stochastically independent. He can observe the value of the target at the same time as finding it and has to immediately decide whether to shoot it or not. If he decides not to shoot, then he comes back home, or else a bullet he fired hits the target with a hit probability $q \in (0,1]$. The game that he bagged can be sold at a price w on his way home. If the bullet does not hit, then the target runs away with a probability $r \in [0,1]$, and he comes back home with no game. When it does not escape, he must decide whether to fire an additional bullet. Assume repeated firings occur at the same point in time.

His purpose is to maximize the total expected reward from the game that will be bagged over a given planning horizon, that is, t periods. Figure 1 illustrates the structure of the decision problem.

Now, since no target being found can be regarded as a target of value 0 being found, the target appearance probability θ and the target value distribution function $F_1(w)$ can be combined into a distribution function $F(w)$ whose probability (density) function is $f(w) = (1-\theta)I(w = 0) + \theta f_1(w)I(w > 0)$ where $f_1(w)$ is a probability (density) function and $I(\mathcal{S}) = 1$ if a given statement $\mathcal{S}$ is true, or else $I(\mathcal{S}) = 0$ †. Let $\mu = \int_0^1 \xi dF(\xi)$. Finally we assume that

$$c \leq \beta q \mu, \tag{1}$$

where $\beta \in (0,1]$ is a discount factor. The assumption implies that it is profitable to go hunting even with one period remaining and only one bullet in hand. Later on in Section 7 we derive the results of the case of $c > \beta q \mu$.

†See preliminaries of [2]

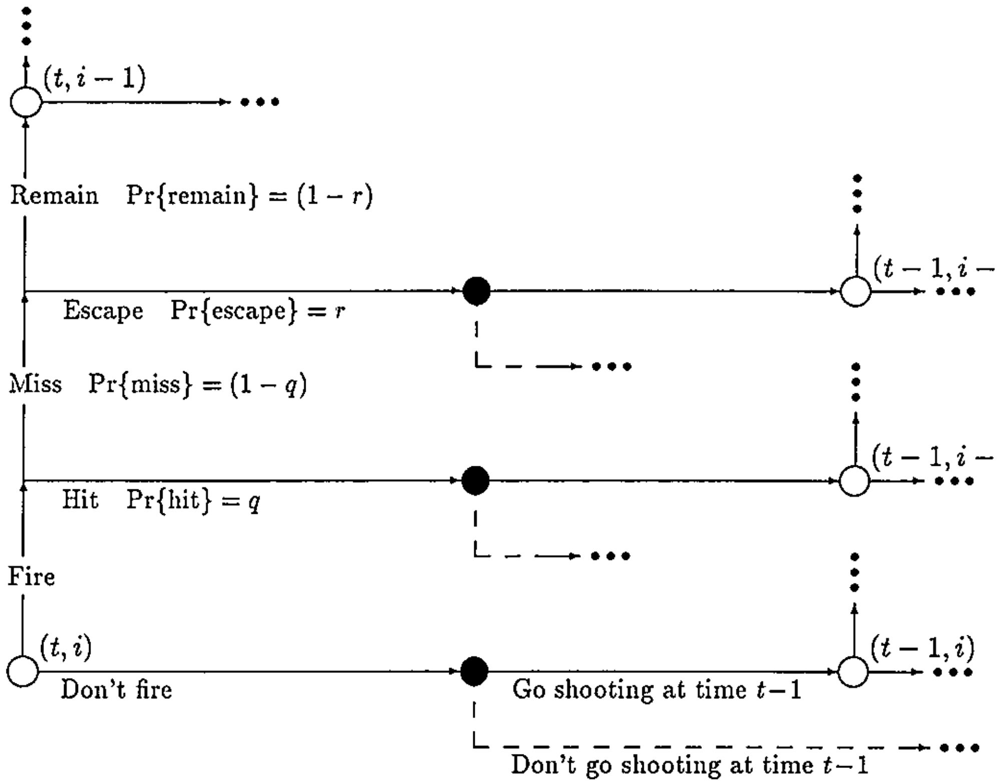


Figure 1 Decision process of this problem

By symbol (t, i), we denote " t periods remaining and i bullets in hand".

3 Fundamental Equations

Let points of time be numbered backward from the final point of the planning horizon as time 0, time 1, and so on; an interval between time t and time $t-1$ is called period t. The objective here is to find an optimal decision policy which maximizes the total expected present discounted reward by allocating the given i bullets to the targets during the given t periods.

Let $u_t(i, w)$ denote the maximum of the total expected present discounted reward when there are t periods remaining and the hunter is seeing a target of value w with i bullets in hand. Furthermore, let $v_t(i)$ denote the expectation of $u_t(i, w)$ in terms of w, that is,

$$v_t(i) = \int_0^1 u_t(i, \xi) dF(\xi), \quad t \geq 0. \tag{2}$$

Then, we have the following recursive relations by the principle of optimality.

$$\begin{aligned} u_t(i, w) &= \max\{z_t(i),\ q(w + z_t(i-1)) + (1-q)(rz_t(i-1) + (1-r)u_t(i-1, w))\} \\ &= \max\{z_t(i),\ pu_t(i-1, w) + qw + (1-p)z_t(i-1)\}, \quad t \geq 1,\ i \geq 1, \end{aligned} \tag{3}$$

where

$$p = (1-q)(1-r) \in [0, 1), \tag{4}$$

$$z_t(i) = \max\{\beta v_{t-1}(i) - c,\ \beta z_{t-1}(i)\}, \quad i \geq 0. \tag{5}$$

Here, p is a probability such that the bullet he fired just before does not hit the target and the target still remains, and $z_t(i)$ implies the maximum of the total expected present discounted reward when there are t periods and i bullets remaining, and he decides not to shoot any more at time t. Furthermore, from the definition of the model, we have the following final conditions:

$$u_t(0,w) = v_t(0) = z_0(i) = z_t(0) = 0, \quad t \geq 0, i \geq 0, \tag{6}$$

$$u_0(i,w) = qw + (1-q)(1-r)u_0(i-1,w) = \frac{1-p^i}{1-p}qw, \quad i \geq 1, \tag{7}$$

$$v_0(i) = \frac{1-p^i}{1-p}q\mu, \quad i \geq 1. \tag{8}$$

Hereafter in this section, we clarify the properties of $u_t(i,w)$ and $v_t(i)$.

Lemma 1

(a) *For any i and any w, both $u_t(i,w)$ and $v_t(i)$ are nondecreasing in t.*

(b) *If $\beta q\mu - c \geq 0$, then $z_t(i) = \beta v_t(i) - c$.*

(c) *When $p > 0$, both $u_t(i,w)$ and $v_t(i)$ are strictly increasing in i for any $t \geq 0$ and $w > 0$. When $p = 0$, they are nondecreasing in i.*

(d) *For any t and any i, $u_t(i,w)$ is nondecreasing in w.*

Proof:

- *Proof of* (a)

  Using (5), we can rewrite (3) as follows. Because $c \leq \beta q\mu$, for any i, we have

$$\begin{aligned} z_1(i) &= \max\{\beta v_0(i) - c,\ \beta z_0(i)\} \\ &= \max\{\frac{1-p^i}{1-p}\beta q\mu - c,\ 0\} \\ &\geq \frac{1-p^i}{1-p}\beta q\mu - c \geq 0 = z_0(i), \end{aligned} \tag{9}$$

hence,

$$\begin{aligned} u_1(i,w) &\geq pu_1(i-1) + qw + (1-p)z_0(i-1) \\ &= pu_1(i-1,w) + qw \\ &\geq p(pu_1(i-2,w) + qw) + qw \\ &\vdots \\ &\geq \frac{1-p^i}{1-p}qw = u_0(i,w). \end{aligned} \tag{10}$$

Immediately from above, we get $v_1(i) \geq v_0(i)$. Now assuming $v_t(i) \geq v_{t-1}(i)$ and $z_t(i) \geq z_{t-1}(i)$ for all $i \geq 0$, we have

$$\begin{aligned} z_{t+1}(i) &= \max\{\beta v_t(i) - c,\ \beta z_t(i)\} \\ &\geq \max\{\beta v_{t-1}(i) - c,\ \beta z_{t-1}(i)\} = z_t(i), \end{aligned} \tag{11}$$

accordingly,

$$\begin{aligned} u_{t+1}(1,w) &= \max\{z_{t+1}(1),\ qw\} \\ &\geq \max\{z_t(1),\ qw\} = u_t(1,w). \end{aligned} \tag{12}$$

Furthermore, assume $u_{t+1}(i-1,w) \geq u_t(i-1,w)$ for any w, then the following can be obtained:

$$\begin{aligned} u_{t+1}(i,w) &= \max\{z_{t+1}(i),\ pu_{t+1}(i-1,w)+qw+(1-p)z_{t+1}(i-1)\} \\ &\geq \max\{z_t(i),\ pu_t(i-1,w)+qw+(1-p)z_t(i-1)\} \\ &= u_t(i,w). \end{aligned} \tag{13}$$

Thus, it follows by double induction that $u_t(i,w)$ is nondecreasing in t for any given i and w, so also are $v_t(i)$ and $z_t(i)$ for any given i.

- *Proof of* (b)

Since $\beta q\mu - c \geq 0$, it is clear that $\beta v_0(i) - c - \beta z_0(i) \geq 0$ for any i. Assume $\beta v_{t-1}(i) - c - \beta z_{t-1}(i) \geq 0$ for any i, that is, $z_t(i) = \beta v_{t-1}(i) - c$. Then, we have

$$\begin{aligned} \beta v_t(i) - c - \beta z_t(i) &= \beta v_t(i) - c - \beta(\beta v_{t-1}(i) - c) \\ &\geq \beta v_{t-1}(i) - c - \beta(\beta v_{t-1}(i) - c) \\ &= (1-\beta)(\beta v_{t-1}(i) - c) \\ &\geq (1-\beta)(\beta v_0(i) - c) \geq 0. \end{aligned} \tag{14}$$

Therefore, we get $z_t(i) = \beta v_{t-1}(i) - c$ for any $t \geq 1$ and $i \geq 1$.

Consequently, it follows for $t \geq 1$ that

$$u_t(i,w) = \begin{cases} \max\{\beta v_{t-1}(i) - c,\ pu_t(i-1,w)+qw+(1-p)(\beta v_{t-1}(i-1)-c)\}, & i \geq 2, \\ \max\{\beta v_{t-1}(1) - c,\ qw\}, & i = 1. \end{cases} \tag{15}$$

- *Proof of* (c)

When $p > 0$, it is obvious from (7) that both $u_0(i,w)$ and $v_0(i)$ are strictly increasing in i. Let $v_{t-1}(i)$ be strictly increasing in i. Then, clearly $u_t(1,w) > u_t(0,w)$ for $w > 0$. Furthermore, we suppose that $u_t(i,w) > u_t(i-1,w)$ for $w > 0$. Then we have

$$\begin{aligned} u_t(i+1,w) &= \max\{\beta v_{t-1}(i+1) - c,\ pu_t(i,w)+qw+(1-p)(\beta v_{t-1}(i)-c)\} \\ &> \max\{\beta v_{t-1}(i) - c,\ pu_t(i-1,w)+qw+(1-p)(\beta v_{t-1}(i-1)-c)\} \\ &= u_t(i,w), \end{aligned} \tag{16}$$

$$\begin{aligned} v_t(i+1) &= \int u_t(i+1,\xi)dF(\xi) \\ &> \int u_t(i,\xi)dF(\xi) = v_t(i). \end{aligned} \tag{17}$$

Hence, if $u_{t-1}(i,w)$ is strictly increasing in i, so also are $u_t(i,w)$ and $v_t(i)$. Thus, it is proven by using induction that $u_t(i,w)$ and $v_t(i)$ are strictly increasing in i for any t and $w > 0$.

For $p = 0$, the proof is basically the same as above.

- *Proof of* (d)

Easily proven by induction. ∎

Using (2) and (15) recursively, we can calculate $v_t(i)$ starting with the final conditions (6), (7) and (8).

4 Structure of Optimal Policy

Now define $g_t(i,w)$ as follows:

$$g_t(i,w) = \begin{cases} pu_t(i-1,w)+qw+(1-p)(\beta v_{t-1}(i-1)-c) - (\beta v_{t-1}(i)-c), & i \geq 2, \\ qw - (\beta v_{t-1}(1)-c), & i = 1. \end{cases} \tag{18}$$

Lemma 2 *For any given $t \geq 1$ and $i \geq 1$, $g_t(i,w) = 0$ has a unique solution $w = h_t(i) \in [0,1)$.*

Proof: It is obvious from Lemma 1(d) that $g_t(i,w)$ is strictly increasing in w. First, for $t \geq 1$, we have

$$g_t(1,0) = -(\beta v_{t-1}(1) - c) \leq -(\beta v_0(1) - c) \leq 0. \tag{19}$$

Suppose $g_t(i-1,0) \leq 0$, that is, $u_t(i-1,0) = \beta v_{t-1}(i-1) - c$. Then, we have

$$\begin{aligned} g_t(i,0) &= pu_t(i-1,0) + (1-p)(\beta v_{t-1}(i-1) - c) - (\beta v_{t-1}(i) - c) \\ &= p(\beta v_{t-1}(i-1) - c) + (1-p)(\beta v_{t-1}(i-1) - c) - (\beta v_{t-1}(i) - c) \\ &= \beta(v_{t-1}(i-1) - v_{t-1}(i)) \leq 0. \end{aligned} \tag{20}$$

Therefore, we obtain $g_t(i,0) \leq 0$ for all $t \geq 1$ and $i \geq 1$.

Next, we shall show $g_t(i,1) > 0$ for $t \geq 1$ and $i \geq 1$. Since there exists a $\xi \in [0,1]$ such that $u_t(i,\xi) < u_t(i,1)$, it follows that

$$v_t(i) = \int_0^1 u_t(i,\xi)dF(\xi) < \int_0^1 u_t(i,1)dF(\xi) = u_t(i,1), \quad i \geq 0. \tag{21}$$

Now, $g_1(1,1) = q(1-\beta\mu) + c > 0$. Let $g_{t-1}(1,1) > 0$, so $u_t(1,1) = q$. Then,

$$g_t(1,1) = q - (\beta v_{t-1}(1) - c) > q - \beta u_{t-1}(1,1) + c = q(1-\beta) + c \geq 0. \tag{22}$$

Therefore, we have $g_t(1,1) > 0$ for all $t \geq 1$. Since $u_1(i-1,1) \geq \beta v_0(i-1) - c$ for $i \geq 2$, we have

$$\begin{aligned} g_1(i,1) &= pu_1(i-1,1) + q + (1-p)(\beta v_0(i-1) - c) - (\beta v_0(i) - c) \\ &\geq q + \beta(v_0(i-1) - v_0(i)) = q(1 - \beta p^{i-1}\mu) > 0. \end{aligned} \tag{23}$$

Suppose $g_{t-1}(i,1) > 0$ for all $i \geq 2$. Then,

$$\begin{aligned} g_t(i,1) &= pu_t(i-1,1) + q + (1-p)(\beta v_{t-1}(i-1) - c) - (\beta v_{t-1}(i) - c) \\ &\geq pu_t(i-1,1) + q + (1-p)(\beta v_{t-1}(i-1) - c) - v_{t-1}(i) \\ &= pu_t(i-1,1) + q + (1-p)(\beta v_{t-1}(i-1) - c) \\ &\qquad - \int \max\{\beta v_{t-2}(i) - c,\ pu_{t-1}(i-1,\xi) + q\xi + (1-p)(\beta v_{t-2}(i-1) - c)\} dF(\xi) \\ &= \int \min\Big\{pu_t(i-1,1) + q + (1-p)(\beta v_{t-2}(i-1) - c) - (\beta v_{t-2}(i) - c), \\ &\qquad\qquad p\Big((u_t(i-1,1) - u_{t-1}(i-1,\xi)\Big) + q(1-\xi)\Big\} dF(\xi) \\ &\geq \int \min\{g_{t-1}(i,1),\ q(1-\xi)\} dF(\xi) > 0. \end{aligned} \tag{24}$$

Hence, we get $g_t(i,1) > 0$ for any $i \geq 1$ and $t \geq 1$. Thus, we verify that $g_t(i,w) = 0$ has a unique solution $w = h_t(i) \in [0,1)$. ∎

We call $h_t(i)$ a critical value when the hunter has i bullets and t periods remaining. In general,

$$\begin{aligned} g_t(i,h_t(i)) &= p\max\{\beta v_{t-1}(i-1) - c,\ u_t(i-1,h_t(i)) + qh_t(i) + (1-p)(\beta v_{t-1}(i-2) - c)\} \\ &\qquad + qh_t(i) + (1-p)(\beta v_{t-1}(i-1) - c) - (\beta v_{t-1}(i) - c) \\ &\geq qh_t(i) + \beta(v_{t-1}(i-1) - v_{t-1}(i)), \end{aligned} \tag{25}$$

accordingly,

$$h_t(i) \leq \beta(v_{t-1}(i) - v_{t-1}(i-1))/q. \tag{26}$$

The above inequalities (25) and (26) are often used in the proofs below.

Lemma 3 *When $p > 0$, for any $i \geq 1$ and $t \geq 1$, the following hold true.*

$$h_t(i) \geq h_t(i+1) \quad \Leftrightarrow \quad h_t(i+1) = \beta(v_{t-1}(i+1) - v_{t-1}(i))/q,$$
$$h_t(i) < h_t(i+1) \quad \Leftrightarrow \quad h_t(i+1) < \beta(v_{t-1}(i+1) - v_{t-1}(i))/q.$$

When $p = 0$, we get $h_t(i+1) = \beta(v_{t-1}(i+1) - v_{t-1}(i))/q$ for any $i \geq 1$ and $t \geq 1$.

Proof: First, we shall verify the case of $p > 0$. If $h_t(i) \geq h_t(i+1)$, then we have

$$g_t(i+1, h_t(i+1)) = qh_t(i+1) + \beta(v_{t-1}(i+1) - v_{t-1}(i)) = 0. \tag{27}$$

Conversely, if $h_t(i+1) = \beta(v_{t-1}(i+1) - v_{t-1}(i))/q$, then we get

$$g_t(i+1, h_t(i+1)) = pu_t(i, h_t(i+1)) - p(\beta v_{t-1}(i) - c) = 0, \tag{28}$$

from which we have

$$u_t(i, h_t(i+1)) = \beta v_{t-1}(i) - c. \tag{29}$$

Immediately, we obtain $h_t(i) < h_t(i+1) \Leftrightarrow h_t(i+1) < \beta(v_{t-1}(i+1) - v_{t-1}(i))/q$ from the above.

When $p = 0$, we can easily verify it because $g_t(i+1, w) = w + \beta(v_{t-1}(i) - v_{t-1}(i+1))$ for all $i \geq 1$ and $t \geq 1$. ∎

Lemma 4 *For $t \geq 1$ and $p > 0$,*

$$\beta(2v_{t-1}(1) - v_{t-1}(2)) - c > (=)\, 0 \;\Leftrightarrow\; h_t(1) > (=)\, h_t(2).$$

Proof: Because $h_t(1) = (\beta v_{t-1}(1) - c)/q$ for $t \geq 1$, we have

$$\begin{aligned} g_t(2, h_t(1)) &= qh_t(1) - \beta(v_{t-1}(2) - v_{t-1}(1)) \\ &= \beta(2v_{t-1}(1) - v_{t-1}(2)) - c, \end{aligned} \tag{30}$$

from which we immediately get the statement in the lemma. ∎

For $t \geq 1$ and $p > 0$, we can easily obtain $\beta(2v_{t-1}(1) - v_{t-1}(2)) - c < 0$ if and only if $h_t(1) < h_t(2)$, as a converse of the above lemma.

Lemma 5 *For a given $i \geq 1$ and $t \geq 1$,*

(a) $h_t(i) > h_t(i+1) \Rightarrow 2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) > 0$,

(b) $h_t(i) \geq h_t(i+1) \Rightarrow 2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) \geq 0$.

Proof: First, we shall prove (a). If $h_t(i) > h_t(i+1)$, then

$$\begin{aligned} 0 < g_t(i+1, h_t(i)) &= qh_t(i) - \beta(v_{t-1}(i+1) - v_{t-1}(i)) \\ &\leq \beta(2v_{t-1}(i) - v_{t-1}(i+1) - v_{t-1}(i-1)), \quad i \geq 2, \end{aligned} \tag{31}$$

$$\begin{aligned} 0 < g_t(2, h_t(1)) &= qh_t(1) - \beta(v_{t-1}(1) - v_{t-1}(2)) \\ &= \beta(2v_{t-1}(1) - v_{t-1}(2)) - c \\ &\leq \beta(2v_{t-1}(1) - v_{t-1}(0) - v_{t-1}(2)). \end{aligned} \tag{32}$$

We can easily show (b) by replacing $<$ with $\leq$ in the proof above. ∎

From Lemma 5(b), we have the following corollary.

Corollary 1 *For a given $i \geq 1$ and $t \geq 1$,*

$$2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) < 0 \;\Rightarrow\; h_t(i) < h_t(i+1).$$

Lemma 6 *If $h_t(i) = \beta(v_{t-1}(i) - v_{t-1}(i-1))/q$ for a given $i \geq 2$, then*

$$2v_{t-1}(i) - v_{t-1}(i+1) - v_{t-1}(i-1) > 0 \;\Rightarrow\; h_t(i) > h_t(i+1).$$

Besides, if $2v_{t-1}(i) - v_{t-1}(i+1) - v_{t-1}(i-1) = 0$, then $h_t(i) \geq h_t(i+1)$.

Proof: Clearly we have

$$\begin{aligned} 0 &< \beta(2v_{t-1}(i) - v_{t-1}(i+1) - v_{t-1}(i-1)) \\ &= qh_t(i) - \beta(v_{t-1}(i+1) - v_{t-1}(i)) \qquad (33) \\ &\leq q(h_t(i) - h_t(i+1)). \qquad (34) \end{aligned}$$

Similarly, we can prove the latter half. ∎

Theorem 1 *The critical value $h_t(i)$ is strictly decreasing in i for a given t if and only if*

$$\beta(2v_{t-1}(1) - v_{t-1}(2)) - c > 0, \qquad (35)$$

and

$$2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) > 0, \qquad (36)$$

for all $i \geq 2$. In addition, $h_t(i)$ is nonincreasing in i if and only if $\beta(2v_{t-1}(1) - v_{t-1}(2)) - c \geq 0$ and $2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) \geq 0$ for all $i \geq 2$.

Proof: Now we shall show the former half of the theorem. If $h_t(i)$ is strictly decreasing in i, then we can obtain

$$\beta(2v_{t-1}(1) - v_{t-1}(2)) - c > 0, \qquad (37)$$

from Lemma 4 and

$$2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) > 0, \qquad (38)$$

for any given $i \geq 1$ by using Lemma 5(a).

The sufficient condition can be proven as follows. From Lemma 4, we have $h_t(1) > h_t(2)$ and

$$h_t(2) = \beta(v_{t-1}(2) - v_{t-1}(1))/q. \qquad (39)$$

By using Lemma 6, we get $h_t(2) > h_t(3)$, so

$$h_t(3) = \beta(v_{t-1}(3) - v_{t-1}(2))/q. \qquad (40)$$

Repeating the same procedure, we obtain $h_t(i) > h_t(i+1)$ for all $i \geq 1$.

In a similar way, we can prove the latter half as in the proof above. ∎

Theorem 2 *If $h_t(i)$ is strictly decreasing (nonincreasing) in i for a given t, then*

$$2v_t(i) - v_t(i-1) - v_t(i+1) > (\geq) 0, \quad i \geq 1.$$

Proof: We shall only prove the case that $h_t(i)$ is strictly decreasing in i. When $h_t(i)$ is nonincreasing, it can be proven by replacing $>$ with $\geq$ in the proof below.

From the hypothesis of the theorem, Lemma 3, and Theorem 1, the following can be said for a given t:

(a) $h_t(i) = \beta(v_{t-1}(i) - v_t(i-1))/q, \quad i \geq 2,$

(b) $\beta(2v_{t-1}(1) - v_{t-1}(2)) - c > 0,$

(c) $2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) > 0, \quad i \geq 2.$

Now, let $A_t(i,\xi)$, $B_t(i,\xi)$ and $C_t(i,\xi)$ satisfy the following equations:

$$2v_t(i) - v_t(i-1) - v_t(i+1) = \int_0^{h_t(i)} A_t(i,\xi)dF(\xi) + \int_{h_t(i)}^{h_t(i-1)} B_t(i,\xi)dF(\xi) + \int_{h_t(i-1)}^{1} C_t(i,\xi)dF(\xi), \quad i \geq 2, \tag{41}$$

$$2v_t(1) - v_t(0) - v_t(2) = \int_0^{h_t(1)} A_t(1,\xi)dF(\xi) + \int_{h_t(1)}^{1} B_t(1,\xi)dF(\xi). \tag{42}$$

First, we have for $0 \leq \xi < h_t(1)$

$$\begin{aligned} A_t(1,\xi) &= 2(\beta v_{t-1}(1) - c) - \max\{\beta(v_{t-1}(2) - c),\ pu_t(1,\xi) + q\xi + (1-p)(\beta v_{t-1}(1) - c)\} \\ &= \min\{\beta(2v_{t-1}(1) - v_{t-1}(2)) - c,\ (\beta v_{t-1}(1) - c) - q\xi\} > 0 \end{aligned} \tag{43}$$

and for $h_t(1) < \xi \leq 1$

$$\begin{aligned} B_t(1,\xi) &= 2q\xi - (1+p)q\xi - (1-p)(\beta v_{t-1}(1) - c) \\ &> (1-p)(qh_t(1) - (\beta v_{t-1}(1) - c)) = 0. \end{aligned} \tag{44}$$

Besides, $A_t(1, h_t(1)) = B_t(1, h_t(1)) = 0$. Because $F(w)$ does not concentrate on a single point, we get $2v_t(1) - v_t(0) - v_t(2) > 0$.

Next, we consider the case of $i \geq 2$. If $0 \leq \xi < h_t(i)$, then

$$\begin{aligned} A_t(i,\xi) &= 2(\beta v_{t-1}(i) - c) - (\beta v_{t-1}(i-1) - c) - \max\{\beta v_{t-1}(i+1) - c,\ q\xi + (\beta v_{t-1}(i) - c)\} \\ &= \min\{\beta(2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1)),\ \beta(v_{t-1}(i) - v_{t-1}(i-1)) - q\xi\} > 0. \end{aligned} \tag{45}$$

If $h_t(i) < \xi \leq h_t(i-1)$, then

$$\begin{aligned} B_t(i,\xi) &= 2(q\xi + (\beta v_{t-1}(i-1) - c) - (\beta v_{t-1}(i-1) - c) \\ &\quad -((1+p)q\xi + (1-p)\beta v_{t-1}(i) + p\beta v_{t-1}(i-1) - c) \\ &= (1-p)\Big(q\xi - \beta(v_{t-1}(i) - v_{t-1}(i-1))\Big) > 0, \end{aligned} \tag{46}$$

If $h_t(i-1) < \xi \leq 1$, then

$$\begin{aligned} C_t(i,\xi) &= 2u_t(i,\xi) - u_t(i-1,\xi) - (pu_t(i,\xi) + q\xi + (1-p)(\beta v_{t-1}(i) - c)) \\ &= -(1-p)^2 u_t(i-1,\xi) + (1-p)q\xi \\ &\quad -\beta(1-p)v_{t-1}(i) + \beta(2-p)(1-p)v_{t-1}(i-1) - (1-p)^2 c \\ &= -(1-p)^2 pu_t(i-2,\xi) + (1-p)pq\xi - \beta(1-p)v_{t-1}(i) \\ &\quad +\beta(2-p)(1-p)v_{t-1}(i-1) - \beta(1-p)^3 v_{t-1}(i-2) - (1-p)^2 pc \\ &= pC_t(i-1,\xi) + \beta(1-p)(2v_{t-1}(i-1) - v_{t-1}(i-2) - v_{t-1}(i)), \quad i \geq 3. \end{aligned} \tag{47}$$

In order to verify $C_t(i,\xi) > 0$ for a given i and $\xi \in (h_t(i-1), 1]$, we must prove $C_t(i-1,\xi) > 0$ for $h_t(i-1) < \xi \leq 1$. Now, we get for $h_t(i-1) < \xi \leq h_t(i-2)$

$$\begin{aligned} C_t(i-1,\xi) &= 2u_t(i-1,\xi) - u_t(i-2,\xi) - (pu_t(i-1,\xi) + q\xi + (1-p)(\beta v_{t-1}(i-1) - c)) \\ &= -(1-p)^2 u_t(i-2,\xi) + (1-p)q\xi \\ &\quad -\beta(1-p)v_{t-1}(i-1) + \beta(2-p)(1-p)v_{t-1}(i-2) - (1-p)^2 c \\ &= (1-p)\Big(q\xi - \beta(v_{t-1}(i-1) - v_{t-1}(i-2))\Big) > 0 \end{aligned} \tag{48}$$

and for $h_t(i-2) < \xi \le 1$

$$C_t(i-1,\xi) = pC_t(i-2,\xi) + \beta(1-p)(2v_{t-1}(i-2) - v_{t-1}(i-3) - v_{t-1}(i-1)), \quad i \ge 4. \tag{49}$$

Therefore, if $C_t(i-2,\xi) > 0$ for $h_t(i-2) < \xi \le 1$, then we get $C_t(i-1,\xi) > 0$ for $h_t(i-2) < \xi \le 1$. Below, it suffices to show $C_t(2,\xi) > 0$ for $h_t(1) < \xi \le 1$ by repeating the same procedure. For $h_t(1) < \xi \le 1$, we obtain

$$\begin{aligned} C_t(2,\xi) &= -(1-p)^2 u_t(1,\xi) + (1-p)q\xi \\ &\quad -\beta(1-p)v_{t-1}(2) + \beta(2-p)(1-p)v_{t-1}(1) - (1-p)^2 c \\ &> (1-p)pqh_t(1) - \beta(1-p)v_{t-1}(2) + \beta(2-p)(1-p)v_{t-1}(1) - (1-p)^2 c \\ &= (1-p)\Big(\beta(2v_{t-1}(1) - v_{t-1}(2)) - c\Big) > 0. \end{aligned} \tag{50}$$

Accordingly, we get $C_t(i,\xi) > 0$ for $h_t(i-1) < \xi \le 1$.

In addition, $A_t(i,h_t(i)) = B_t(i,h_t(i)) = 0$. Because $F(w)$ does not concentrate on a single point, we get $2v_t(i) - v_t(i-1) - v_t(i+1) > 0$ for a given $i \ge 2$. Eventually, it follows that $2v_t(i) - v_t(i-1) - v_t(i+1) > 0$ for all $i \ge 1$ if $h_t(i)$ is strictly decreasing in i. ∎

5 Case with No Search Cost, i.e. $c = 0$

Lemma 7 *If $p > 0$, then $h_t(i)$ is strictly decreasing in i for all $t \ge 1$, and if $p = 0$, then nonincreasing in i.*

Proof: First, we shall prove the case of $p > 0$. Then, it is obvious that

$$2v_0(i) - v_0(i-1) - v_0(i+1) > 0, \tag{51}$$

for all $i \ge 1$. From Lemma 4, $h_1(2) < h_1(1)$, so $h_1(2) = \beta(v_0(2) - v_0(1))/q$. By using Lemma 6 repeatedly, we can show that $h_1(i)$ is strictly decreasing in i. From Theorem 2, we get

$$2v_1(i) - v_1(i-1) - v_1(i+1) > 0, \tag{52}$$

for all $i \ge 1$. Repeating this procedure in t, we obtain that $h_t(i)$ is strictly decreasing in i for any given $t \ge 1$.

When $p = 0$, it is clear that $2v_0(i) - v_0(i-1) - v_0(i+1) \ge 0$ for all $i \ge 1$, so $h_1(i)$ is nonincreasing in i. From Theorem 2, $2v_1(i) - v_1(i-1) - v_1(i+1) \ge 0$ for all $i \ge 1$. Repeating this procedure, we can prove it. ∎

The following corollary is clear from Theorem 2.

Corollary 2 *When $p > 0$, $v_t(i)$ is strictly concave in i for all $t \ge 0$. When $p = 0$, it is concave in i.*

Furthermore, if $\beta = 1$, then we can get the following statement:

Corollary 3 *If $\beta = 1$ and $c = 0$, then $h_t(i)$ is nondecreasing in t for any given $i \ge 1$.*

Proof: It is clear that $u_t(1,w) - u_t(0,w) \ge v_{t-1}(1) - v_{t-1}(0)$. Suppose

$$u_t(i-1,w) - u_t(i-2,w) \ge v_{t-1}(i-1) - v_{t-1}(i-2), \quad i \ge 1. \tag{53}$$

Then, we obtain for $0 \le w \le h_t(i)$

$$u_t(i,w) - u_t(i-1,w) = v_{t-1}(i) - v_{t-1}(i-1), \tag{54}$$

for $h_t(i) < w \le h_t(i-1)$

$$\begin{aligned} u_t(i,w) - u_t(i-1,w) &= u_t(i,w) - v_{t-1}(i-1) \\ &> v_{t-1}(i) - v_{t-1}(i-1) \end{aligned} \tag{55}$$

and for $h_t(i-1) < w \le 1$

$$\begin{aligned} u_t(i,w) - u_t(i-1,w) &= p(u_t(i-1,w) - u_t(i-2,w)) + (1-p)(v_{t-1}(i-1) - v_{t-1}(i-2)) \\ &\ge v_{t-1}(i-1) - v_{t-1}(i-2) && (56) \\ &> v_{t-1}(i) - v_{t-1}(i-1). && (57) \end{aligned}$$

In transforming (56) into (57), we used Corollary 2. The above relation still holds even if we integrate (57) in terms of w. Hence, we have

$$v_t(i) - v_t(i-1) \ge v_{t-1}(i) - v_{t-1}(i-1), \quad i \ge 1. \tag{58}$$

Since $h_t(i)$ is nonincreasing in i for all $t \ge 1$, $h_t(i)$ always equals to $\beta(v_{t-1}(i) - v_{t-1}(i-1)/q$. Therefore, $h_t(i)$ is nondecreasing in t for all $i \ge 1$. ∎

6 Case with Search Cost, i.e. $c > 0$

We can conjecture that the more bullets the hunter may have, he may shoot at even the smaller valued targets, that is, $h_t(i)$ is decreasing in i. However $h_t(i)$ is not always decreasing in i if $c > 0$. Now we shall show a simple example where $h_t(i)$ is not decreasing in i.

Because $h_1(1) = (\beta v_0(1) - c)/q$, we have

$$\begin{aligned} g_1(2, h_1(1)) &= \max\{qh_1(1) + \beta(v_0(1) - v_0(2)), \\ &\qquad (1+p)qh_1(1) + (1-p)(\beta v_0(1) - c) - (\beta v_0(2) - c)\} \\ &= \beta(2v_0(1) - v_0(2)) - c \\ &= \beta(1-p)q\mu - c. \end{aligned} \tag{59}$$

If $\beta(1-p)q\mu < c \le \beta q\mu$, then $g_1(2, h_1(1)) < 0$. Therefore, in this case, it follows that $h_1(1) < h_1(2)$. Like this, it may be possible that $h_t(i)$ becomes increasing in i for a certain interval. To put it more concretely, $h_t(i)$ may become such as that depicted in Figure 2. The figure implies the following in terms of the optimal decision policy.

(1) Suppose a present target value is w_a. If he has more than eleven bullets, then he should continue to fire until either; he gets it, the remaining number of bullets becomes less than twelve or it escapes. If he starts with less than twelve, then he should not fire and search for the next target.

(2) Suppose a present target value is w_b. If he has more than six bullets, then he should continue to fire until either; he gets it, the number of bullets in hand becomes less than seven or it runs away. If he starts with more than two and less than seven, then he had better not fire and search for the next target. If he has one or two bullets, then he should fire until either; he gets it, spends all the bullets or it escapes. Here, it should be noted that there may exist two critical points i_* and i^* ($i_* < i^*$) in terms of the number of bullets for one target value (think about critical points of i for w_b). It goes without saying that if $h_t(i)$ is nonincreasing in i, then such a thing never occurs.

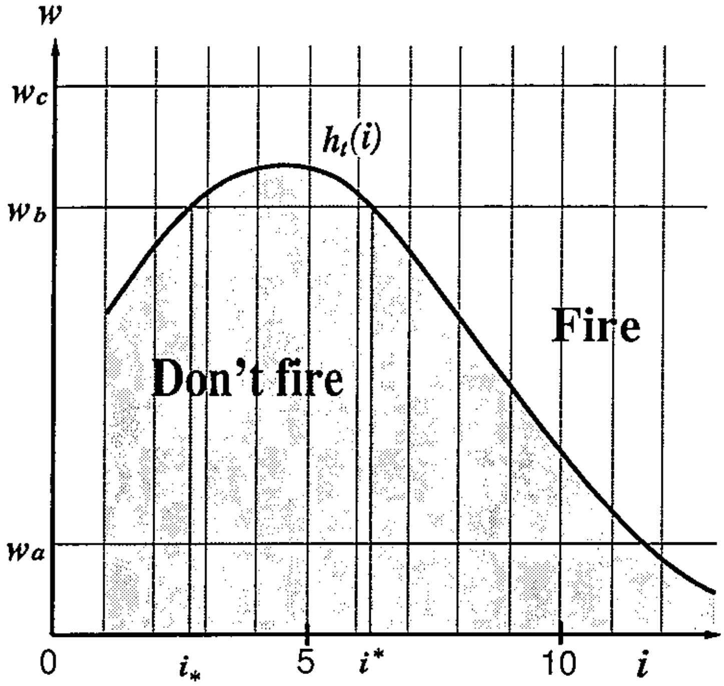


Figure 2 Relation between i and w

(3) Suppose $w = w_c$. In this case, he should continue to fire until either; he gets it, spends all the bullets or it runs away.

Now we have $\lim_{i\to\infty} h_t(i) = 0$ because $\beta(v_{t-1}(i) - v_t(i-1))/q$ converges to zero as $i \to \infty$. Therefore, this leads us to the following lemma.

Lemma 8 *If $h_t(i_a) < h_t(i_a + 1)$ for a certain i_a, then there exists $i_b > i_a$ such that $h_t(i_b) > h_t(i_b + 1)$.*

On the other hand, we get the following condition by Lemma 4, Lemma 5, and Theorem 2.

Condition 1 *A necessary and sufficient condition for which $h_t(i)$ is nonincreasing in i for all $t \geq 1$ is*

$$\beta(2v_t(1) - v_t(2)) - c \geq 0, \tag{60}$$

for all $t \geq 0$.

From the above condition, we shall show the existence of intervals of c, denoted by Lemma 9 that will be proven below. In preparation of the description and proof of Lemma 9, we define some symbols as follows. Because we can regard $v_t(i)$ as a function of c, we shall use the symbol $v_t(i, c)$ instead of $v_t(i)$ if necessary. In the same way, $h_t(i, c)$ will be also used. Let

$$D_t(c) = \beta(2v_t(1, c) - v_t(2, c)) - c. \tag{61}$$

Let the limits of $v_t(i, c)$, $h_t(i, c)$ and $D_t(c)$ as $t \to \infty$ be designated by the symbols without the subscript t, i.e. $v(i, c)$, $h(i, c)$ and $D(c)$.

Lemma 9 *If $\beta < 1$, then there exists a positive number c_* for which $h_t(i, c)$ is nonincreasing in i for any t and $c \in [0, c_*]$. In addition, if $\beta < 1$ and $p > 0$, then there exists $c^* \in [c_*, \beta q\mu)$ for which $h_t(1) < h_t(2)$ for any t and $c \in [c^*, \beta q\mu]$.*

Proof: Throughout the proof, we discuss only the case of $\beta < 1$. It is clear that $v_t(1,c)$ and $v_t(2,c)$ are continuous functions of $c \in [0, \beta q\mu]$ for any given $t \geq 0$. Now we have

$$\begin{aligned} v(1,c) - v_t(1,c) &= \int \max\{\beta v(1,c) - c, q\xi\} dF(\xi) - \int \max\{\beta v_{t-1}(1,c) - c, q\xi\} dF(\xi) \\ &\leq \int \max\{\beta(v(1,c) - v_{t-1}(1,c)), 0\} dF(\xi) \\ &= \beta(v(1,c) - v_{t-1}(1,c)) \\ &\vdots \\ &\leq \beta^t(v(1,c) - v_0(1,c)) < \beta^t. \end{aligned} \tag{62}$$

Let $T = \log\varepsilon / \log\beta$. Then, for any $t > T$, we get

$$0 \leq v(1,0) - v_t(1,0) < \beta^t < \varepsilon. \tag{63}$$

Therefore, $v_t(1,c)$ uniformly converges to $v(1,c)$ as $t \to \infty$, so $v(1,c)$ is also a continuous function of c. In a similar way, $v(2,c)$ also can be shown to be a continuous function of c. Eventually, it follows that $D_t(c)$ is a continuous function of c for any given t, and so also is $D(c)$.

If $c = 0$, then $h_t(1,0) \geq h_t(2,0)$ for all $t \geq 1$ from Lemma 7. And we get $h_t(2,0) < 1$ for all t from Lemma 2. Hence we have

$$\begin{aligned} D_t(0)/\beta &= 2v_t(1,0) - v_t(2,0) \\ &= \int_0^{h_t(2,0)} \beta(2v_{t-1}(1,0) - v_{t-1}(2,0)) dF(\xi) \\ &\quad + \int_{h_t(2,0)}^{h_t(1,0)} (\beta v_{t-1}(1,0) - q\xi) dF(\xi) + \int_{h_t(1,0)}^1 (1-p)(q\xi - \beta v_{t-1}(1,0)) dF(\xi) \\ &= q\int_0^{h_t(2,0)} (h_t(1,0) - h_t(2,0)) dF(\xi) \\ &\quad + q\int_{h_t(2,0)}^{h_t(1,0)} (v_t(1,0) - \xi) dF(\xi) + q\int_{h_t(1,0)}^1 (1-p)(\xi - v_t(1,0)) dF(\xi) > 0. \end{aligned} \tag{64}$$

In addition, we get $D(0) > 0$ since $h(1,0) = \beta v(1,0)/q \leq \beta < 1$.

On the other hand, if $p > 0$, then we have

$$v(1,\beta q\mu) = v_t(1,\beta q\mu) = q\mu, \tag{65}$$

$$v(2,\beta q\mu) \geq v_t(2,\beta q\mu) = \int \max\{\beta v_{t-1}(2,\beta q\mu) - \beta q\mu,\ (1+p)q\xi\} dF(w) \geq (1+p)q\mu, \tag{66}$$

therefore,

$$D(\beta q\mu) \leq D_t(\beta q\mu) \leq 2\beta q\mu - \beta(1+p)q\mu - \beta q\mu < 0. \tag{67}$$

If $p = 0$, then we obtain

$$v_t(1,\beta q\mu) = q\mu, \tag{68}$$

$$v_t(2,\beta q\mu) = \int \max\{\beta v_{t-1}(2,\beta q\mu) - \beta q\mu,\ q\xi\} dF(w) = q\mu \tag{69}$$

where the third term of (69) can be easily proven by induction. Consequently, we get

$$D_t(\beta q\mu) = 2\beta q\mu - \beta q\mu - \beta q\mu = 0. \tag{70}$$

Because $v_t(1,\beta q\mu) = v_t(2,\beta q\mu) = q\mu$ for all $t \geq 0$ and $p = 0$, we have $v(1,\beta q\mu) = v(2,\beta q\mu) = q\mu$, so $D(\beta q\mu) = 0$. Accordingly, we come to:

(a) Let $c_{t*} = \min\{c \mid D_t(c) = 0,\ c \in [0, \beta q\mu]\}$. Then, $c_{t*} > 0$ for all t.

(b) Let $c_t^* = \max\{c \mid D_t(c) = 0,\ c \in [0, \beta q\mu]\}$ and $p > 0$. Then, $c_t^* \in [c_{t*}, \beta q\mu)$ for all t.

Furthermore, define $c_* = \inf_{t \geq 0} c_{t*}$ and $c^* = \sup_{t \geq 0} c_t^*$. Then, we have $0 < c_* \leq c^* \leq \beta q\mu$ for any p since $D(0) > 0$, and $c^* < \beta q\mu$ for $p > 0$ since $D(\beta q\mu) < 0$. Therefore, the length of the interval $[0, c_*]$ is not zero for any q and r, and the length of the interval $[c^*, \beta q\mu]$ is also not zero for $p > 0$. With this the proof is complete. ∎

Remark When $\beta = 1$, we have $v(1,0) \geq q\xi$ for all $\xi \in [0,1]$ since $v(1,0) = \int \max\{v(1,0), q\xi\} dF(\xi)$ and $F(\xi) < 1$ for $\xi < 1$. Hence, $v(1,0) = q$. And we have $v(2,0) = 2q$, so, $D(0) = 0$.

7 Case with a Sufficiently Large Search Cost, i.e. $c > \beta q\mu$

Now, using (2), (3) and (5), we shall examine the case of $c > \beta q\mu$.

$$\begin{aligned} u_t(i,w) &= \max\{z_t(i),\ pu_t(i-1,w) + qw + (1-p)z_t(i-1)\},\ t \geq 1,\ i \geq 1, \\ v_t(i) &= \int_0^1 u_t(i,\xi) dF(\xi)\ , \\ z_t(i) &= \max\{\beta v_{t-1}(i) - c,\ \beta z_{t-1}(i)\},\ t \geq 1. \end{aligned}$$

Suppose $c \geq \beta(1-p^i)q\mu/(1-p)$. Then $\beta v_0(i) - c \leq 0$ for all i, so

$$z_1(i) = \max\{\beta v_0(i) - c,\ \beta z_0(i)\} = 0. \tag{71}$$

Hence, we get

$$\begin{aligned} u_1(i,\xi) &= \max\{z_1(i), pu_t(i-1,\xi) + q\xi + (1-p)z_1(i-1)\} \\ &= p\max\{z_1(i-1), pu_t(i-2,\xi) + q\xi + (1-p)z_1(i-2)\} + q\xi \\ &\ \ \vdots \\ &= \frac{1-p^i}{1-p} q\xi. \end{aligned} \tag{72}$$

Accordingly, we obtain

$$\begin{aligned} z_2(i) &= \max\{\beta \int u_1(i,\xi) dF(\xi) - c, \beta z_1(i)\} \\ &= \max\{\beta \frac{1-p^i}{1-p} q\mu - c, 0\} = 0, \\ &\ \ \vdots \\ z_t(i) &= 0. \end{aligned} \tag{73}$$

As a result, in this case, an optimal policy is not to go shooting at all.

Next, suppose $\beta q\mu < c < \beta(1-p^i)q\mu/(1-p)$. Now let $c = \beta(1-p^{\kappa(c)})q\mu/(1-p)$, from which we have

$$\kappa(c) = \log_p\Big(1 - \frac{(1-p)c}{\beta q\mu}\Big). \tag{74}$$

Then, for $i \geq \kappa(c)$, we get

$$\begin{aligned} z_1(i) &= \max\{\beta v_0(i) - c,\ \beta z_0(i)\} = \beta v_0(i) - c, \\ z_2(i) &= \max\{\beta v_1(i) - c,\ \beta z_1(i)\} = \beta v_1(i) - c, \\ &\ \ \vdots \\ z_t(i) &= \beta v_{t-1}(i) - c. \end{aligned} \tag{75}$$

On the other hand, if the number of bullets becomes less than $\kappa(c)$, then it follows from the same reason above that $z_t(i) = 0$.

Hence, in this case, the optimal decision policy becomes identical to the case of $c \le \beta q\mu$ till the remaining number of bullets becomes less than $\kappa(c)$. If the number becomes less than $\kappa(c)$, it is optimal not to go shooting from the next time point on. However, if he has not got the present target yet and it still remains, then he should continue to fire until either; he gets it, spends all the bullets or it runs away.

8 Case of Infinite Bullets

Since both $u_t(i,w)$ and $v_t(i)$ are increasing in i and upper-bounded for any $t \ge 0$, they have finite limits as $i \to \infty$ for any given t. Let $u_t(w) = \lim_{i\to\infty} u_t(i,w)$ and $v_t = \lim_{i\to\infty} v_t(i)$. Then, immediately we get $u_0(w) = qw/(1-p)$ and $v_0 = q\mu/(1-p)$. By induction, it can be easily shown for any given $t > 1$,

$$u_t(w) = \begin{cases} \dfrac{qw}{1-p} + t(\dfrac{q\mu}{1-p} - c), & \beta = 1, \\ \dfrac{qw}{1-p} + \dfrac{1-\beta^t}{1-\beta}(\dfrac{\beta q\mu}{1-p} - c), & \beta < 1, \end{cases} \tag{76}$$

$$v_t = \begin{cases} \dfrac{(t+1)q\mu}{1-p} - tc, & \beta = 1, \\ \dfrac{1-\beta^{t+1}}{1-\beta}\dfrac{q\mu}{1-p} + \dfrac{1-\beta^t}{1-\beta}c, & \beta < 1, \end{cases} \tag{77}$$

from which we have

$$u(w) = \lim_{t\to\infty} u_t(w) = \begin{cases} \infty, & \beta = 1, \\ \dfrac{qw}{1-p} + \dfrac{1}{1-\beta}(\dfrac{\beta q\mu}{1-p} - c), & \beta < 1, \end{cases} \tag{78}$$

$$v = \lim_{t\to\infty} v_t = \begin{cases} \infty, & \beta = 1, \\ \dfrac{1}{1-\beta}(\dfrac{\beta q\mu}{1-p} - c), & \beta < 1. \end{cases} \tag{79}$$

In addition, we have already mentioned the following in Section 6.

$$h_t = \lim_{i\to\infty} h_t(i) = 0, \quad t \ge 1. \tag{80}$$

9 Results of Numerical Examples

In this section, by using numerical examples, we will examine the relationship of $h_t(i)$ with t, i, and other related parameters, β, q, r, and c. Here, we use a discrete uniform distribution function with 101 mass points, equally spaced on $[0, 1]$. All the figures below illustrate the relation between i and $h_t(i)$. Now, we summarize the implications that are deduced from Figures 3 – 8.

(1) If $c = 0$, then the critical value $h_t(i)$ is nonincreasing in i for all $t \ge 1$ (See Figure 3(a,d)).

(2) If $c > 0$, then $h_t(i)$ is not always decreasing in i (See Figure 3(b,c,e)). We see that the position of the maximal value shifts to the right as the planning horizon becomes longer. But the maximal value does not always exist. In fact, as seen in Figure 3(c), it doesn't appear if the planning horizon is less than or equal to five.

Table 1: Values of $D_t(c)$ (q=0.9, r=0.1)

t	$\beta = 0.9$, $c = 10^{-2}$	$\beta = 1$, $c = 10^{-2}$	$\beta = 1$, $c = 10^{-4}$	$\beta = 1$, $c = 10^{-5}$
0	0.36755000	0.40850000	0.40949000	0.40949900
1	0.19188098	0.21291981	0.21395330	0.21396271
2	0.15190272	0.16463288	0.16586718	0.16587843
3	0.13178598	0.13661378	0.13805264	0.13806577
4	0.12041073	0.11747495	0.11911669	0.11913169
5	0.11373292	0.10330834	0.10515039	0.10516724
10	0.10474813	0.06461562	0.06742139	0.06744734
50	0.10410564	0.00998829	0.01818530	0.01828057
100	0.10410564	0.00015102	0.00946883	0.00964782
500	0.10410564	-0.00194995	0.00126448	0.00194734
1000	0.10410564	-0.00194995	0.00007398	0.00085488
1500	0.10410564	-0.00194995	-0.00015226	0.00045219
2000	0.10410564	-0.00194995	-0.00018908	0.00024057
2500	0.10410564	-0.00194995	-0.00019427	0.00011673
3000	0.10410564	-0.00194995	-0.00019494	0.00004239
3500	0.10410564	-0.00194995	-0.00019502	-0.00000193
4000	0.10410564	-0.00194995	-0.00019503	-0.00002805
4500	0.10410564	-0.00194995	-0.00019503	-0.00004304
5000	0.10410564	-0.00194995	-0.00019503	-0.00005149

(3) In Figure 3(f), $h_t(i)$ is nonincreasing in i in spite of c being positive. From Condition 1, $h_t(i)$ is nonincreasing in i for all $t \geq 1$ if and only if $D_t(c) \geq 0$ for all $t \geq 0$. The results of the calculation of $D_t(c)$ for $t = 0, 1, \ldots, 5000$ are as shown in Table 1 [‡]. In the table, $D_t(c)$ seems to converge to a positive number $0.104\cdots$ for $\beta = 0.9$ and $c = 10^{-2}$. However it becomes negative for large t when $\beta = 1$ and $c = 10^{-2}$. Even if $\beta = 1$ and $c = 10^{-5}$, $D_{3500}(10^{-5})$ is negative. From these, we can expect that $h_t(i)$ is nonincreasing in i for all $t \geq 1$ if and only if $c = 0$ when $\beta = 1$ except the case of $p = 0$.

(4) If $q = 1$, then $h_t(i)$ is nonincreasing in i for any set of β and c (See Figure 4). This case is reduced to an optimal stopping problem in which i offers can be accepted. We shall conjecture that the property holds for $r = 1$.

(5) We can regard $h_t(i)$ as a function of β, q and r similar to c. So we can use $h_t(i, \beta, q, r, c)$ instead of $h_t(i)$. The critical value $h_t(i, \beta, q, r, c)$ is nondecreasing in β and nonincreasing in r and c for given other parameters (See Figure 5, 7, and 8). These properties, if they are true, don't conflict with our intuition.

(6) From Figure 6, we find that $h_t(i, \beta, q, r, c)$ is not monotone in q for other given parameters. It may be natural that the fact is concerned with the property such that $h_t(i)$ is not always decreasing in i.

(7) From Figure 3–8, we can guess that the critical value $h_t(i)$ is nondecreasing in t for all $i \geq 1$ not only in the case of $\beta = 1$ and $c = 0$, but also in the case of any set of β and c.

[‡]The distribution used in the example is a continuous uniform one on [0,1].

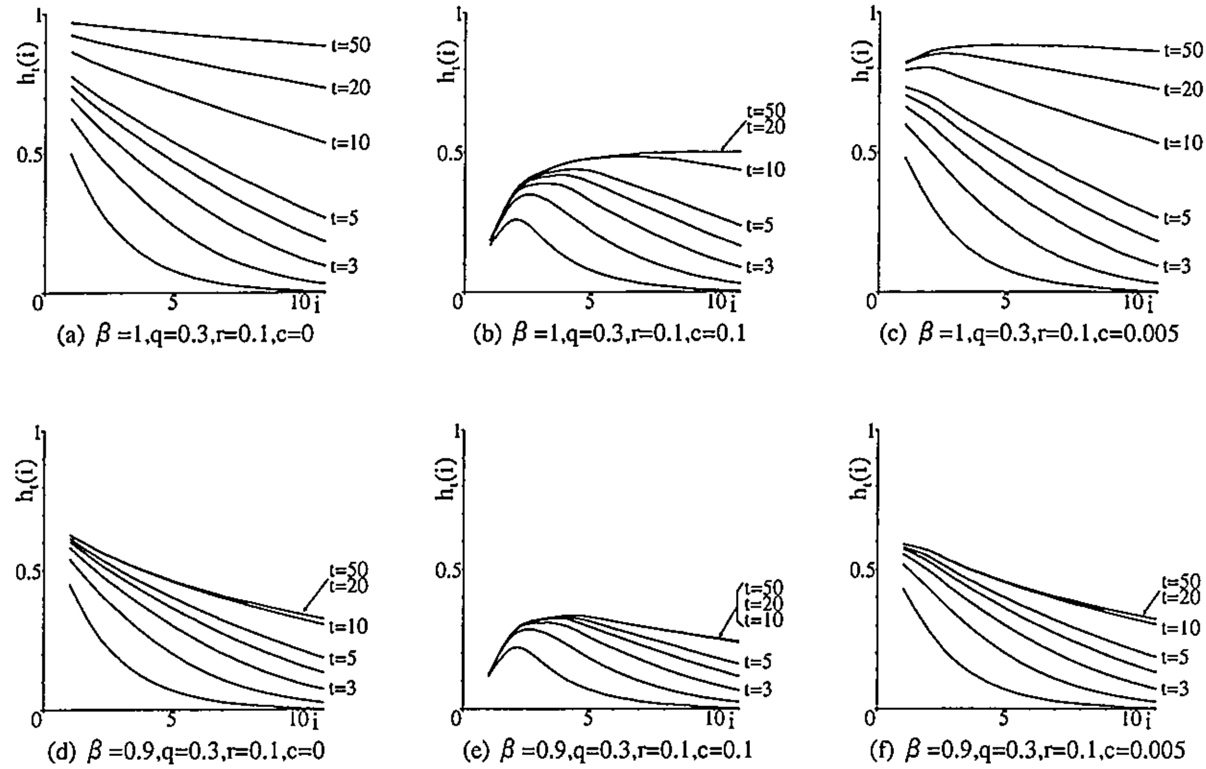


Figure 3 Relations between i and $h_t(i)$

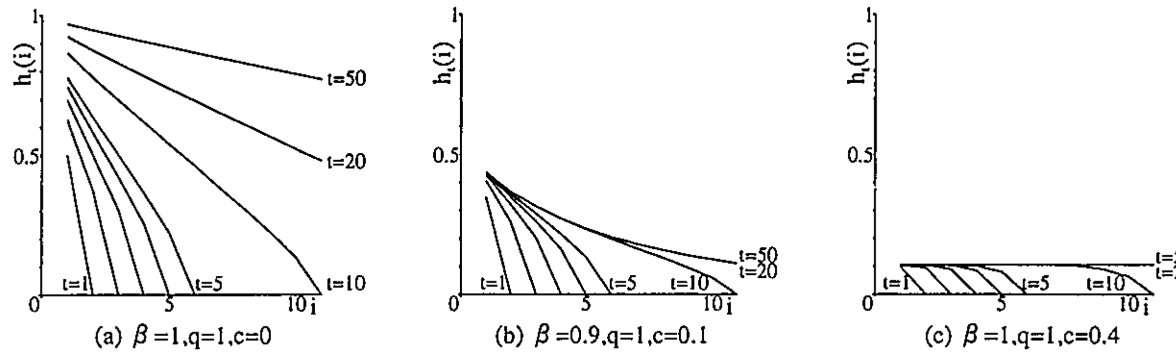
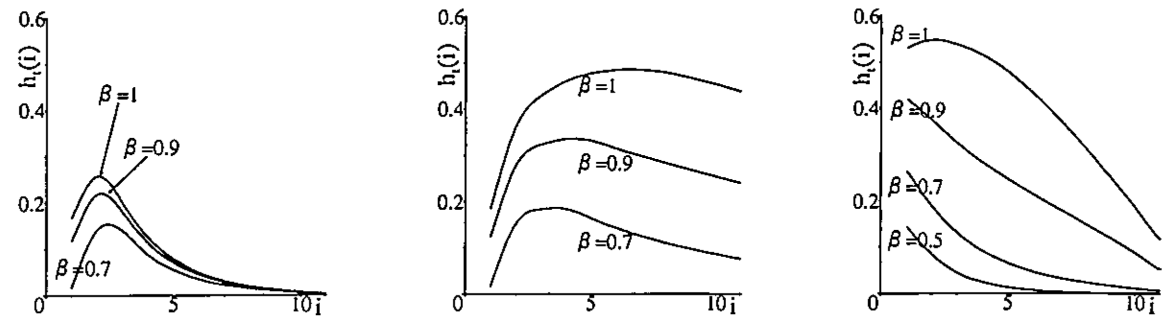
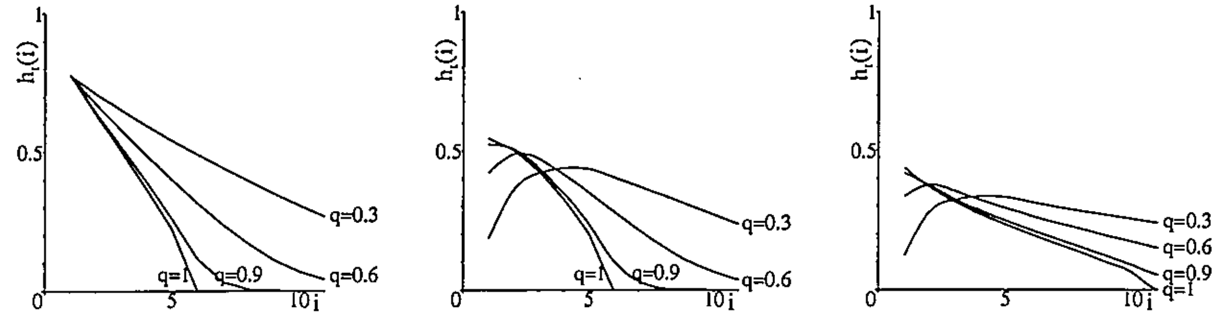


Figure 4 Relations between i and $h_t(i)$ for $q = 1$



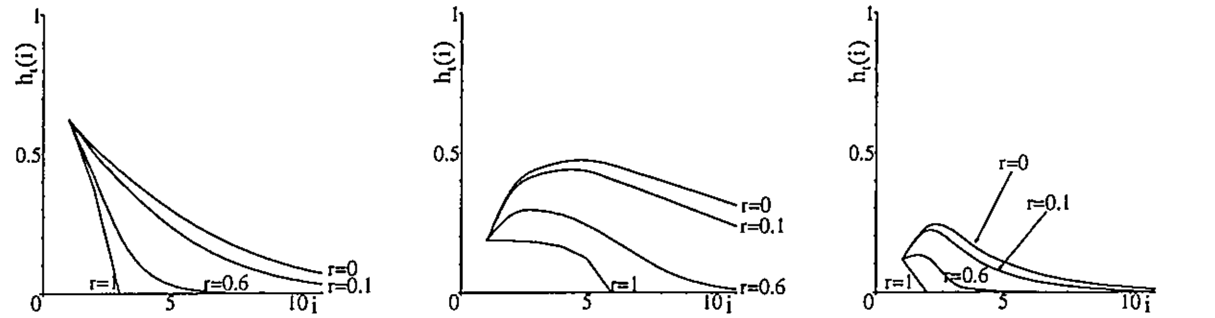
(a) q=0.3,r=0.1,c=0.1,t=1 (b) q=0.3,r=0.1,c=0.1,t=10 (c) q=0.9,r=0.1,c=0.1,t=10

Figure 5 Sensitivity of β



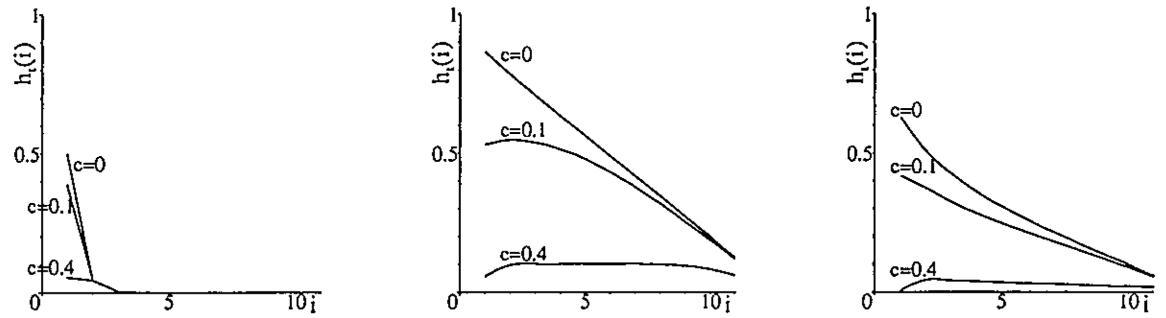
(a) β =1,r=0.1,c=0,t=5 (b) β =1,r=0.1,c=0.1,t=5 (c) β =0.9,r=0.1,c=0.1,t=10

Figure 6 Sensitivity of q



(a) β =1,q=0.3,c=0,t=2 (b) β =1,q=0.3,c=0.1,t=5 (c) β =0.9,q=0.3,c=0.1,t=1

Figure 7 Sensitivity of r



(a) β =1,q=0.9,r=0.1,t=1 (b) β =1,q=0.9,r=0.1,t=10 (b) β =0.9,q=0.9,r=0.1,t=10

Figure 8 Sensitivity of c

10 Conclusions

- *Optimal Decision Policy*

First, if $c \leq \beta q\mu$, then the policy is as follows. If you have at least one bullet, then go shooting. Suppose you see a target whose value is w when you have i bullets and t periods remaining. In this case, the following can be said:

(1) If $w < h_t(i)$, then search for the next target by paying c.

(2) If $w \geq h_t(i)$, $w \geq h_t(i-1), \ldots, w \geq h_t(i-k)$ and $w < h_t(i-k-1)$, $0 \leq k \leq i-1$, then continue to fire until either; you get the target, the number of remaining bullets becomes less than k or it runs away.

(3) If $w \geq h_t(k)$ for all $k = 1, 2, \ldots, i$, then fire until either; you get the target, spend all the bullets or it escapes.

Next, we summarize the policy in the case for $c > \beta q\mu$. When you have i bullets in hand and t periods remain, the optimal policy is as follows:

(1) If $i \leq \kappa(c)$, then don't go shooting over the whole planning horizon.

(2) If $i > \kappa(c)$, then go shooting till the number of bullets on hand becomes less than $\kappa(c)$, and the optimal decision policy for firing is similar to the case of $c \leq \beta q\mu$. Suppose the number has become less than $\kappa(c)$. Then, you should not go shooting from the next time point on. However, if you have not got the present target yet and it still remains, continue to fire until either; you get it, spend all the bullets or it escapes.

- *Properties of the Critical Value*

We have obtained that, if $c = 0$, then $h_t(i)$ is nonincreasing in i for all $t \geq 1$. On the other hand, if $c > 0$, then we get the following results:

(1) If $h_t(i_a) < h_t(i_a + 1)$ for a certain i_a, then there exists $i_b > i_a$ such that $h_t(i_b) > h_t(i_b + 1)$.

(2) If $\beta < 1$, then $h_t(i, c)$ is nonincreasing in i for all t and $c \in [0, c_*]$ where $c_* > 0$.

(3) If $\beta < 1$ and $p > 0$, then $h_t(i, c)$ is not nonincreasing in i for any $t \geq 1$ and $c \in [c^*, \beta q\mu]$ with $c^* \in [c_*, \beta q\mu)$, that is, $h_t(i, c)$ always has an interval of i on which it strictly increases.

11 Some Limitations

We have discussed a sequential allocation problem with search cost where a shoot-look-shoot policy is employed. However it is sure that some of the assumptions restrict the problem. We will suggest the following provisions to relax these restrictions: (1) The hit probability q depends on the total number of bullets he has shot at a present target. This can be also said for the probability of escaping, r. (2) In this paper, the action of shooting results in only one of two outcomes; "get the target" or "don't get the target". As a variation of this problem, we can consider the case where a piece of a target can be got in exchange for a bullet. Then, there still remains a decision whether to shoot the rest of the target or not. (3) When he is going to chase a target, it may happen that he finds a new target, that is, he can find more than two targets in one period. (4) He can replenish some bullets by paying some cost if the number of remaining bullets becomes below a certain level.

REFERENCES

[1] C. Derman, G. J. Lieberman and S. M. Ross, A Stochastic Sequential Allocation Model, *Operations Research*, Vol.23, 1120-1130(1975)

[2] S. Ikuta, Optimal Stopping Problem with Uncertain Recall, *Journal of the Operations Research Society of Japan*, Vol.31, 145-170(1988)

[3] S. Ikuta, The Optimal Stopping Problem in which the Sum of the Accepted Offer's Value and the Remaining Search Budget is an Objective Function, *Journal of the Operations Research Society of Japan*, Vol.35, 172-193(1992)

[4] T. Kisi, Suboptimal Decision Rule for Attacking Targets of Opportunity, *Naval Research Logistics Quarterly*, Vol.23, 525-533(1976)

[5] D. V. Mastran and C. J. Thomas, Decision Rule for Attacking Targets of Opportunity, *Naval Research Logistics Quarterly*, Vol.20, 661-672(1973)

[6] T. Namekata, Y. Tabata and T. Nishida, A Sequential Allocation Problem with two kinds of Targets, *Journal of the Operations Research Society of Japan*, Vol.22, 16-28(1979)

[7] G. P. Prastacos, Optimal Sequential Investment Decisions under Conditions of Uncertainty, *Management Science*, Vol.29, 118-134(1983)

[8] M. Sakaguchi, A Sequential Allocation Problem for Randomly Appearing Targets, *Mathematica Japonica*, Vol.21, 89-103(1976)