DISCUSSION PAPER NO. 316

# OPTIMAL SEARCH OVER SETS OF DISTRIBUTIONS

by

Daniel F. Spulber

November 1977

This research was supported by The National Science Foundation
Grant No. SOC76-20953

Summary

This paper examines optimal strategies for a number of search problems where the decision maker randomly selects an element of a set of investment opportunities each yielding an unknown return. Simple stopping rules are characterized for the cases where the next available opportunity is and is not known at the time the decision is made whether to stop or continue.

## Introduction

Many sequential decision problems in economics can be analysed as search problems where there are a number of different investment opportunities each involving an independent probability distribution of possible rewards. Weitzman [5] and Burdett [1] consider search problems where a decision maker may select from a set of these distributions. In some economic applications however, opportunities may not all be available at the same time but may present themselves randomly to the decision maker. We will consider the search problem encountered by a decision maker who has the option of either searching further the currently available opportunity (and foregoing all others) or observing the next available opportunity. We will consider two variations of this search problem. The first case will require that the searcher decide whether to stop or continue before finding out the next available opportunity and the second case will allow the searcher to know the next opportunity. We will show that, for a decision maker maximizing the net expected (discounted) reward, the optimal search strategy will involve simple myopic decision rules. These rules will involve a partition of the set of opportunities based on the implicit worth of pursuing each opportunity.

Assumptions and Definitions

The type of problem we will deal with in this paper can be considered within the category of box games. (This term applies to various search models such as [1] and [5]). We will take box games to mean the problem of sampling at a constant cost from a box where the rewards from sampling are uncertain but the distribution of rewards is known. We thus allow the possibility of different rewards from a given box.

The standard search problem involves repeated sampling from a single box where the cost of an additional sample is some constant $c$, the reward from a sample $X$ has the known distribution $F(\cdot)$. When recall is allowed, the strategy which maximizes the expected value of search is to continue sampling if and only if the best reward received is less than $\omega$ where $\omega$ is given by:

$$c = \beta\int_{\omega}^{\infty}(X - \omega)dF(X) \quad + \quad (\beta-1)\omega$$

where $\beta$ is the discount rate. The 'reservation wage' $\omega$ will not depend on the number of samples available to the searcher (see [1], [5]). If recall is not allowed, a constant reservation rule will be optimal when the searcher is permitted an infinite number of samples.

We will consider a search problem where a decision maker must draw a box at random from a set of boxes of various types. The first part of the paper will consider various games when the searcher does not know what box will be available next after the currently available box is sampled. The second part will consider the possibility that the type of box that will be available next is known. This distinction will have an important effect on the optimal decision rules.

We will now describe a game with a set J containing K types of boxes, $J = \{1, \ldots, K\}$. We will use the following assumptions throughout the paper:

A1. Each type of box $j \in J$ has a known independent distribution $Fj(\cdot)$ of rewards $yj$ obtained by opening a box of type j. Also $E(y_j) < \infty$ for all j.

A2. The available type of box for each period is determined by a random draw from the set J. The probability of drawing any box j is given by the known probabilities $Pj$ where $0 \leq Pj \leq 1$, $j = 1, \ldots, K$ and $\Sigma_{j=1}^{k} P_j = 1$.

A3. Continuing the game for another period entails a fixed cost c. Net rewards are discounted at rate $\beta$.

A4. The best reward observed may be recalled.

A5. The searcher has the option of not drawing any boxes from J and searching further the currently available box; i.e., the decision maker has three options.
   i) Accept the last reward and stop.
   ii) Search further the last box which was drawn from J (and lose the option of selecting another box from J).
   iii) Continue search by selecting a box at random from J.

We will now consider two alternative additional assumptions. The first of these will be considered in this section of the paper, the second in the next section.

A6a. The searcher simultaneously draws a box from the set J and takes a single sample from that box.

A6b. The searcher samples the currently available box and selects the next available box from J although that box is not sampled unless the choice is made to continue sampling from J.

To understand the effect of these last two assumptions it is interesting to note that if A5 were replaced by the assumption:

A5'. The searcher may only take a single sample from an available box before drawing another box from J.

then two well known models would be obtained. Assumptions A1 to A4, A5' and A6a would be the standard single box search problem where the distribution of rewards is a mixture of the K distributions. Assumptions A1 to A4, A5' and A6b will yield a special case of a search problem considered by Derman [ 2 ], Lippman and McCall [ 3 ] and others who allow the probability of the next box type selected to depend on the box currently available (i.e., to obtain this special case let $P_{ij} = P_j$ for all $i, j \in J$). Lippman and McCall interpret the available box as the state of the economy and the distribution of rewards from any given box as the distribution of available wage offers. Derman considers the problem of periodic inspection of a system or its components where the available box is the current state of the unit in question.

Thus, the most interesting assumption here will be A5 which allows the option of searching further the currently available box. The force of this assumption will be seen to be a simple decision rule in terms of a partition of the set J based on the implicit worth of each type of box j as defined in the single box game.

There are various economic applications of the model developed here. Within a job search model, the various types of boxes can be interpreted as different signals obtained through investment in human capital. Thus, a sample from J will involve continuing education for an additional period to observe the signal which will determine the worker's type. There will be different wage offer distributions associated with each worker type. The workers' options, assuming that wage offers are received while in school, will be to accept the last wage offer, or drop out to search (and lose the option of further altering the signal) or to continue receiving wage offers and continue education.

Another framework to which the model presented here applies is the research and development problem. Let each type of box represent a state of a research process and the distribution of rewards for each box represent the possible developed innovations resulting from a particular state of the research process. The decision maker's options, assuming that the development process is carried on simultaneously with research will be to install the last innovation developed, or to halt the research process and seek the best developed innovation for the last state of research or to continue both the research and development processes.

Game 1 The next box available is not known.

We are now ready to consider the game as defined by A1 to A5 and A6a. The strategy that maximizes the expected value of the game will not depend on the number of possible periods remaining in the game. Let $\omega j$ represent the stopping reward if box $j$ is used in the single box game, i.e.

$$\omega j = -c + \beta\int_{\omega j}^{\infty} (yj - \omega j)dFj(yj) + \beta\omega j$$

where $y_j$ is the expected reward from opening a box of type $j$. Let $y$ be the best reward received to date. We can let $J = \{1, . . ., K\}$ represent the types of boxes contained in the set J. Renumber the set J so that $\omega_j \leq \omega_{j+1}$, $j = 1, . . . , K - 1$. Consider an arbitrary partition of $J$ given by $J' = \{1, . . ., K'\}$ and $J'' = \{K' + 1, . . . K\}$.

Proposition 1.

There exists a unique value $z$ of drawing another box from J which induces a partition of $J$ such that $z > \omega_j$ for $j \in J'$ and $z \leq \omega_j$ for $j \in J''$. The value of $z$ is given by

$$z = \Sigma_{j \in J'} P_j \; [\beta \int_z^\infty (y_j - z) dFj(y_j) + \beta z - c] + \Sigma_{j \in J''} P j \omega j \quad .$$

The strategy which maximizes the value of the game depends only on the best reward received to date $y$ and on the stopping reward of the box $i$ which has just been opened:

i) If $i \in J'$, the draw again from J if $y < z$, otherwise stop and receive $y$.

ii) If $i \in J''$, do not draw again from $J$ (i.e., continue single box game for $i \in J''$).

Proof.

The value of continuing is the same regardless of the state or the number of steps remaining. Let this value be represented by $z$. The value of the game from the current date onwards, if the best reward to date is $y$ and the last box opened is $i$, is given by:

$$V(i,y) = \max \{y, \omega_i, z\}.$$

The value of continuity is defined by

$$-c + \beta \Sigma_{j=1}^{K} P_j \int_0^\infty V(j, y_j) dF_j(yj).$$

If $z$ is the rule implicit in the value function $V(j, y_j), \forall_j$, we will show that there exists a unique $z$ which solves:

$$z = H(z)$$

where $H(\cdot)$ is given by

$$\begin{aligned}
H(z) &= -c + \beta \Sigma_j^K P_j \int_0^\infty V(j, y_j) dFj(yj) \\
&= -c + \beta \Sigma_j^K P_j \int_0^\infty \max\{y_j, \omega_j, z\} dF_j(yj) \\
&= -c + \beta \Sigma_{j=1}^{K'(z)} P_j \left[\int_z^\infty (y_j - z) dF_j(yj) + z\right] \\
&\quad + \beta \Sigma_{j=K'(z)+1}^{K} P_j \left[\int_{\omega_j}^\infty (y_j - \omega_j) dF_j(yj) + \omega_j\right] \\
&= \Sigma_{j=1}^{K'(z)} P_j \left[\beta \int_z^\infty (y_j - z) dF_j(yj) + \beta z - c\right] \\
&\quad + \Sigma_{j=K'(z)+1}^{K} P_j \left[\beta \int_{\omega_j}^\infty (y_j - \omega_j) dF_j(yj) + \beta \omega_j - c\right] \\
&= \Sigma_{j=1}^{K'(z)} P_j \left[\beta \int_z^\infty (yj - z) dF_j(yj) + \beta z - c\right] + \Sigma_{j=K'(z)+1}^{K} P_j \omega_j \quad .
\end{aligned}$$

So, $H(z) = \Sigma_{j=1}^{K'(z)} P_j H_j(z) + \Sigma_{j=K'(z)+1}^{K} P_j \omega_j$

where $H_j(z) = [\beta \int_z^\infty (y_j - z) dF_j(yj) + \beta z - c]$.

We know that the $H_j$ are continuous and monotonic. To prove the existence and uniqueness of $z$ defined by the equation

$$H(z) = z$$

it remains to show that $H$ is continuous and that monotonity holds for $(H(z) - z)$.

For $z \leq \omega_1$ , $H(z) = \Sigma_{j=1}^{K} P_j \omega_j$ .

For $z > \omega_k$ , $H(z) = \Sigma_{j=1}^{K} P_j H_j(z)$ $(K'=K)$

Also for $K'$ fixed, i.e., for $z \in (\omega_j, \omega_{j+1}]$ ,

$$H(z) = \Sigma_{j=1}^{K'} P_j H_j(z) + \Sigma_{j=K'+1}^{K} P_j \omega_j$$

so that $H$ is continuous on the intervals $(\omega_j, \omega_{j+1})$.

It remains to show that $H$ is continuous at each $\omega_j$. Note that

$$H(\omega_i) = \Sigma_{j=1}^{i-1} P_j H_j(\omega_i) + \Sigma_{j=i}^{K} P_j \omega_j \quad .$$

For $z \in (\omega_{i-1}, \omega_i]$, $K' = i - 1$ so that $\lim_{z \to \omega_i} H(z) \to H(\omega_i)$ by the continuity of the $H_j(\cdot)$ $j=1, \ldots K$.

For $z \in (\omega_i, \omega_{i+1}]$, $K' = i$, $H(z) = \Sigma_{j=1}^{i} P_j H_j(z) + \Sigma_{j=i+1}^{K} P_j \omega_j$ .

So for $z \in (\omega_i, \omega_{i+1}]$,

$$H(\omega_i) - H(z) = [\Sigma_{j=1}^{i-1} P_j H_j(\omega_i) + \Sigma_{j=1}^{K} P_j \omega_j] - [\Sigma_{j=1}^{i} P_j H_j(z) + \Sigma_{j=i+1}^{K} P_j \omega_j]$$

$$= \Sigma_{j=1}^{i-1} P_j (H_j(\omega_i) - H_j(z)) + P_i(\omega_i - H_i(z)) \quad .$$

So $\lim_{z \to \omega_i^+} [H(\omega_i) - H(z)]$

$$= \lim_{z \to \omega_i^+} [\Sigma_{j=1}^{i-1} P_j (H_j(\omega_i) - H_j(z))] + \lim_{z \to \omega_i^+} [P_i(\omega_i - H_i(z))]$$

First, $\lim_{z \to \omega_i^+} [\Sigma_{j=1}^{i-1} P_j (H_j(\omega_i) - H_j(z))] = 0$ by the continuity of

$H_j(\cdot)$ for $j = 1, \ldots, K$.

Secondly, $\lim_{z\to\omega_i^+} P_i(\omega_i - H_i(z)) = 0$ by the continuity of the $H_j(\cdot)$ and the fact that $H_i(\omega_i) = \omega_i$ (by definition).

Clearly, $H(z) - z$ is continuous. Also,

$$(H(z) - z) = \Sigma_{j=1}^{K'(z)} P_j(H_j(z) - z) + \Sigma_{j=K'(z)+1}^{K} P_j(\omega_j - z)$$ is monotonic

so that a solution exists to the equation $0 = (H(z) - z)$ which is unique.

Clearly, if all of the $F_j(\cdot)$ have the same mean and are numbered in order of increasing risk (in the sense of Rothschild-Stiglitz; see [4]), then the stopping rewards $\omega_j$, $j=1, \ldots, K$ will be ordered in terms of increasing size. If the game is stopped for $j \in J''$ it will be stopped for $j + 1$. This implies the following result:

Proposition 2. If $F_i(\cdot)$ and $F_j(\cdot)$ have the same mean, and $F_i(\cdot)$ is riskier than $F_j(\cdot)$ (in the sense of Rothschild-Stiglitz) then if $j \in J''$, $i \in J''$.

Proof. This follows from the fact that $j \in J''$ implies $\omega_j \geq z$. Since $\omega_i \geq \omega_j$ (by the definition of Rothschild-Stiglitz) this implies $\omega_i \geq z$, so $i \in J''$.

Game 2. The next box available is known.

Consider the game defined by A1 to A5 and A6b. The searcher will observe a sample reward from the currently available box and will also select a box at random from the set J. The searcher will then decide whether to accept the best observed reward, or to search the currently available box further or to continue the game by sampling the next available box and drawing again from J. We obtain the following result:

Proposition 3. The stopping rule does not depend on the number of periods remaining. It only depends on the best reward observed to date and on the currently held but unopened box.

The value of searching J further is given by

$$z_i = -c + \beta \Sigma_{j=1}^{K'} P_j H_i(\omega_j) + \beta \Sigma_{j=K'+1}^{K} P_j H_i(z_j)$$

where $H_i(x) = [\int_x^\infty (y - x) dF_i(y) + x]$ .

For this game a partition of J is induced where $w_j \geq z_j$ for $j \in J' = \{1, . . ., K'\}$, and $\omega_j < z_j$ for $j \in J'' = \{K' + 1, . . ., K\}$.
If $j \in J'$ do not search J further (Apply single box game rule.)
If $j \in J''$, stop if $y \geq z_j$ and search J if $y < z_j$ .

Proof:

If $x$ is the best reward observed to date and $i$ is the (unopened) box last drawn from $J$ then the value of state $(i,x)$ is given by

$$V(i,x) = \max\{x,\omega_i, - c + \beta[ \Sigma_{j=1}^K P_j\int_o^\infty V(j,y)dF_i(y)\}$$

where $\omega_i$ is the standard stopping reward for the single box game given by:

$$\omega_i = -c + \beta\, \omega_i + \beta \int_{\omega_i}^\infty (X - \omega_i)dF_i(X).$$

In what follows let $H_i(z) \equiv \int_z^\infty (X - z)dF_i(z) + z.$

If the value of continuing to search $J$ (and taking only one sample from the currently held box $i$) is given by $z_i$ then

$$V(i,x) = \max \{x,\omega_i,z_i\}$$

where $z_i$ solves

$$z_i = -c + \beta [ \Sigma_{j=1}^K P_j \int_o^\infty \max \{y,\omega_j,z_j\}dF_i(y)]$$

$$= -c + \beta\, \Sigma_j^K P_j[ \int_{z_j}^\infty \max \{y,\omega_j\}\ dF_i(y) + \int_o^{z_j} \max \{\omega_j,z_j\}dF_i(y)]$$

$$= -c + \beta\, \Sigma_j^K\ P_j[ \int_{z_j}^\infty \max \{y,\omega_j\}dF_i(y) + \max \{\omega_j,z_j\} \int_o^{z_j}dF_i(y)] \quad .$$

Suppose that $\{z_1, . . ., z_K\}$ are well defined. Partition $J$ such that for $\omega_j \geq z_j$ , $j \in J' = \{1, . . ., K'\}$ and for $\omega_j < z_j$, $j \in J'' = \{K' + 1, . . ., K\}$ (where $J$ has been renumbered appropriately).

Then

$$z_i = -c + \beta\Sigma_{j=1}^{K'} P_j[\int_{\omega_j}^{\infty} (y - \omega_j)dF_i(y) + \omega_j]$$

$$+ \beta\Sigma_{j=K'+1}^{K} P_j[\int_{z_j}^{\infty} (y - z_j)\, dF_i(y) + z_j]$$

$$= -c + \beta\Sigma_{j=1}^{K'} P_jH_i(\omega_j) + \beta\Sigma_{j=K'+1}^{K} P_jH_i(z_j)$$

$$i = 1, \ldots, K \quad .$$

It now remains to show the existence of the $z_i$ . This involves examining the solution to the system of equations given by:

$$z_i = -c + \beta\Sigma_{j=1}^{K} P_j[\int_{z_j}^{\infty} \max\{y,\omega_j\}\, dF_i(y) + \max\{\omega_j, z\}\int_{o}^{z_j} d\,F_i(y)]$$

$$i = 1, \ldots, K \quad .$$

Note that since the expected rewards from sampling a given box once will be bounded (by <u>A1</u>), the implicit worth of a single box, thus the implicit worth of continuing, will be bounded. Thus the function

$$G(z_1, \ldots, z_k) = (G_1(z_1, \ldots, z_k), \ldots, G_k(z_1, \ldots, z_k))$$

where

$$G_i(z_1, \ldots, z_k) \equiv \beta\, \Sigma_{j=1}^{K} P_j \int_{z_j}^{\infty} \max\{y,\omega_j\}dF_i(y) + \max\{\omega_j, z_j\}\int_{o}^{z_j} dF_i(y)]$$

$$i = 1, \ldots K$$

will map a nonempty, compact, convex set into itself and will have a fixed point if each $G_i$ is continuous (by Brouwer's theorem). To show that $G_i$ is continuous in $(z_1, \ldots, z_K)$ for all $i$ look at each element in the sum:

$$\int_{z_j}^{\infty} \max\{y,\omega_j\} dF_i(y) + \max\{\omega_j, z_j\} \int_0^{z_j} dF_i(y) \quad .$$

The first term is continuous since

$$\int_{z_j}^{\infty} \max\{y,\omega_j\} dF_i(y) = \begin{cases} \omega_j(F_i(\omega_j) - F_i(z_j)) + \int_{\omega_j}^{\infty} y\, dF_i(y) & \text{for } z_j \leq \omega_j \\ \int_{z_j}^{\infty} y dF_i(y) & \text{for } z_j \geq \omega_j \end{cases}$$

The second term is continuous since the product of continuous functions is continuous. Note also that each $G_i(z_1, \ldots, z_n)$ is monotonic in its arguments.

Note that the value of continuing will increase with the riskiness of the distribution of the next available box as in the single box game.

<u>Proposition 4</u>. If the means of $F_i(\cdot)$ and $F_\ell(\cdot)$ are equal and $F_i(\cdot)$ is riskier than $F_\ell(\cdot)$ in the sense of Rothschild-Stiglitz then $z_i > z_\ell$.

<u>Proof</u>.

$$z_i = -c + E(F_i(y)) + \beta \Sigma_{j=1}^{K'} P_j [\int_0^{\omega_j} F_i(y)dy] + \beta \Sigma_{j=K'+1}^{K} P_j [\int_0^{z_j} F_i(y)dy]$$

$$z_i - z_\ell = \beta \Sigma_{j=1}^{K'} P_j \int_0^{\omega_j} (F_i - F_\ell)dy + \beta \Sigma_{j=K'+1}^{K} P_j \int_0^{z_j} (F_i - F_\ell)dy \quad .$$

So $(z_i - z_\ell) > 0$ by the Rothschild-Stiglitz definition.

Note that Proposition 3 still holds if the probability of choosing any particular box on the next draw is conditionally dependent on the type of box currently held. If this assumption is made, then game 2 is simply the Lippman and McCall model with the added option of further searching any available distribution of rewards. Proposition 4 will be altered unless assumptions are made on the conditional probabilities.

## Conclusion

We have shown the existence of simple optimal stopping rules when a searcher faces a set of opportunities each yielding an uncertain return and must sample randomly from that set. Of particular interest is the result that when the searcher is allowed to sample further the box currently available to him, a partition of the set of types of boxes is induced. The searcher will apply the single box search rule for a subset of the set of boxes J and will only consider another draw from J on the complement of that subset. This result holds regardless of whether or not the searcher knows which box will be available to him next when he makes his decision as to whether or not to continue to sample from the set of boxes.

References

1. Burdett, Kenneth; "Box Games",University of Wisconsin, unpublished, October 1977.

2. Derman, Cyrus; "On optimal replacement rules when changes of state are Markovian" in Mathematical Optimization Techniques (R. Bellman, Ed.), University of California Press, Berkeley, 1963.

3. Lippman, Steven A., and McCall, John J.; "Job Search in a Dynamic Economy," J.E.T., vol. 12, no. 3, June 1976, pp. 365-390.

4. Rothschild, M., and Stiglitz, J.E.; "Increasing Risk, I: A Definition," J.E.T., vol. 2 (1970), pp. 225-243.

5. Weitzman, Martin L.; "A Basic Result in the Theory of Optimal Search" Massachusetts Institute of Technology, unpublished, July 1977.