DISCUSSION PAPER NO. 353

Optimal Policies and Steady-State Solutions for Inventory Problems with Markovian Uncertainty

by

David Easley* and Daniel F. Spulber**

June 1978
Revised January 1979

* Northwestern University

** Brown University

This research was supported by National Science Foundation Grant Number SOC 76-20953 -- Principle Investigator - John O. Ledyard.

Summary

The existence and optimality of a state-dependent (s,S) policy is shown for the infinite horizon inventory control problem with general Markovian disturbances using Markov contraction operators. A new result is then presented concerning the convergence to steady-state probability distributions of Markov processes in general metric spaces having regeneration points. The optimal stock level for the inventory problem is then shown to be a regeneration point of the inventory process, yielding convergence to a steady-state probability distribution on stock levels.

Introduction

The class of stock adjustment problems known as inventory control has found wide application in economics and operations research (see Arrow, Karlin and Scarf [2], Arrow, Karlin and Suppes [3], and Scarf, Gilford and Shelly [20]). Specific applications and extensions include minimizing the costs of holding and altering inventories of goods when faced with penalties for beginning a period with insufficient stock, management of a money - bond portfolio when transactions costs are present, investment in a capital stock with costs of adjustment, and management of a renewable resource with harvesting costs. In these and other applications, the nature of environmental and market uncertainty may be quite complex. That is, the random disturbances faced by the decision-maker may take values in a general (uncountable) state space and may be generated by Markov processes. This is a reasonable assertion in the case of random demands for the inventory problem or random interest rates in the portfolio problem. Thus, the standard assumptions that disturbances take values in countable state spaces and are independently and identically distributed are approximations at best (see for example [2], [11], [12] and [22]).

This paper will extend the analysis of the inventory problem to the case of Markov disturbances taking values in $\mathbb{R}^n$. An (s,S) type feedback policy, which depends upon the last observed disturbance, will be obtained by using Markov contraction operators on Banach spaces. The optimal policy thus obtained will then be used to demonstrate convergence to an equilibrium probability distribution on stock levels. To demonstrate convergence a renewal theorem of some independent interest will be stated and proved.

Part I of the paper will obtain the existence and optimality of a state-dependent (s,S) inventory control policy using the contraction operator approach to stochastic dynamic programming developed by Blackwell [6], Maitra [17], Strauch [21] and others (see Hinderer [10] and Bertsekas [5] for a discussion). The contraction operator approach will also allow a clearer presentation of the existence and optimality of the (s,S) policy for the special case of independent and identically distributed disturbances.

The optimal stock-adjustment rules obtained from inventory-type problems are useful in examining the long-run behavior of the system since they have the property that when they are applied, the stochastic process describing the changes in the stock over time may enter a 'regeneration state.' This 'regeneration state' is an initial starting point for the random movement of stock levels over time which allows us to characterize the long-run behavior of the stock in an interesting way. It is then possible to show that when a stochastic process obtained from the interaction of an optimal policy and exogenous uncertainty has a regeneration point, the sequence of probabilities generated by the process will converge to an equilibrium (or invariant) distribution over stocks. This invariant distribution of stocks characterizes both the empirical distribution over time for a stock managed by a single decision maker as well as the cross-sectional distribution of stocks for a group of independent and identical decision makers.

The optimal inventory policy guarantees that the regeneration point can be reached in a finite number of periods from any point in the state space. A theorem on ergodicity of stochastic processes will be presented in Section II. This theorem shows that any stochastically bounded process with a regeneration state having a positive probability of recurrence in a finite number of steps, will converge to a unique invariant probability. While this theorem has formal similarities to well known theorems in renewal theory, the result is presented in a more direct manner which will greatly facilitate its application to a wide range of economic problems.

The problem of convergence to an equilibrium probability distribution has been previously considered in the literature on inventory control. We briefly consider three important papers. Karlin [13] pp. 234-237 and [14] shows convergence to an equilibrium distribution for the simple inventory model subject to independent and identically distributed disturbances by directly examining the transition operator for the Markov process of stock levels in [13], and also by renewal theory Methods in [14]. These proofs do not apply when exogenous disturbances are Markovian. Furthermore, it may not be evident how to represent changes in stock levels by sums of i.i.d. random variables when the transition equation is nonlinear. The case of Markov disturbances in an inventory model with a finite state space was considered in Karlin and Fabens [15]. However an (s,S) policy was arbitrarily imposed. In the present paper, the (s,S) policy is derived from optimizing behavior and shown to depend upon the last disturbance when disturbances are Markov. This is a case which cannot be handled by the approach of Karlin and Fabens [15].

In Section III, the optimal policy obtained in Section I, for the case of general Markov disturbances, will be shown to imply the existence of a regeneration point which can be reached from any point in the state space in a finite number of steps. In particular, the optimal stock level for an (s,S)-type stock adjustment policy will be identical to the regeneration point. The Regeneration Theorem of Section II will then be used to show that the stochastic process of stock levels converges to an equilibrium probability on states of the system. Specific examples of economic problems will then be interpreted.

I. The Optimal Inventory-Control Policy

An extension of the Arrow, Harris and Marshak [1] inventory problem will now be presented. The model will be extended to the case of Markov disturbances taking values in a general state space. In addition, a general transition equation will be introduced which yields the inventory control problem as a special case. A decision-maker is assumed to manage a stock so as to minimize expected discounted costs over an infinite horizon. The problem will be reduced to the recursive form of the Bellman optimality equation in which the decision maker applies a stationary-feedback policy. The existence and optimality of this policy will be demonstrated using the techniques of stochastic dynamic programming.

Let x_t be the initial stock at the beginning of period t. The initial stock may be negative indicating shortages or debt. Let $w_t \in E \subset \mathbb{R}^n$ be the random disturbance which is observed during period t. Given the state of the system at time t (s_t, w_t) the transition equation of the system g will describe the remaining stock z_t at the time the decision is made to reorder:

$$z_t = g(x_t, w_t) \tag{1.1}$$

Let a_t be the choice of the initial stock for the next period:

$$x_{t+1} = a_t \quad . \tag{1.2}$$

The stock is assumed to take values in the compact set $[\underline{x}, \overline{x}]$. The bounds may be derived from the costs of holding large inventories and of experiencing large shortages, or from institutional constraints. The following assumption on g is not overly restrictive:

<u>Assumption 1</u> The transition equation g is given by a function $\bar{g}$ which is linear and non-decreasing in x and continuous in w, where

$$g(x,w) = \begin{cases} \bar{x} & \bar{g}(x,w) \geq \bar{x} \\ \bar{g}(x,w) & \text{otherwise} \\ \underline{x} & \bar{g}(x,w) \leq \underline{x} \end{cases}$$

This assumption permits the standard inventory transition equation

$$z_t = x_t - w_t$$

where x_t is the initial stock and w_t represents random demand. For the portfolio problem $\bar{g}$ may take the form

$$z_t = w_t^1 x_t + w_t^2$$

where w_t^1 is a random interest rate applied to savings x_t and w_t^2 represents random endowments or expenditures. Aggregate investment problems may also be handled. Let the capital stock z_t be given by

$$z_t = f(w_t)x_t$$

where x_t is the capital-labor ratio and $f(w_t)$ represents the effects of population growth and capital depreciation. For the investment problem at the level of the firm, $f(w_t)$ may represent the effects of random usage rates and depreciation on the initial capital stock x_t.

The stock is subject to random disturbances which are represented by a Markov process $w \in E \subset \mathbb{R}^n$ with transition probability $\varphi(\cdot|w)$.

<u>Assumption 2</u> The measure φ takes E into the space of probability measure on E, $P(E)$, continuously when $P(E)$ is endowed with the weak topology.

The decision maker will face holding and shortage costs $h(z_t)$ of holding stock z_t for period t.

Assumption 3. The cost of holding and shortage $h(\cdot)$ is continuous, bounded and convex for $z \in [\underline{x},\overline{x}]$.

See Scarf [19] p. 197 for a discussion of these costs for the inventory problem. When the addition to stock $(a_t - z_t)$ is positive, the decision maker will face a 'set-up' cost of ordering K as well as linear ordering costs $c \cdot (a - z)$. The fixed cost K may be due to transactions costs or adjustment costs. The cost function faced by the manager is therefore

$$C(z,a) = h(z) + \delta_{(a-z)} \cdot K + c \cdot (a - z) \qquad 1.3$$

where δ is an indicator function. Given these assumptions, the decision maker's problem is to minimize the expected value of costs discounted at rate α, where $0 < \alpha < 1$. Formally the decision maker solves:

$$\min E \left[\Sigma_{t=1}^{\infty} \alpha^{t-1} C(z_t, a_t) \right] \qquad 1.4$$

$$\text{s.t. } z_1, w_1 \text{ given.}$$

$$z_t \leq a_t \leq \overline{x}$$

$$z_{t+1} = g(a_t, w_{t+1})$$

$$x_{t+1} = a_t$$

The problem may be rewritten in the recursive form of dynamic programming, with value function V defined by

$$V(x,w) = \min_{a \in [z,\bar{x}]} [h(z) + \delta_{(a-z)} \cdot K + c \cdot (a-z) + \alpha \int V(a,w')d\varphi(w'/w)] \qquad 1.5$$

where $z = g(x,w)$ is the remaining stock when the decision is made to reorder. The decision-maker should add to his inventory if and only if there exists $a \in (z,\bar{x}]$ such that:

$$K + c\cdot(a-z) + \alpha\int V(a,w')\,d\varphi(w'/w) < \alpha\int V(z,w')\,d\varphi(w'/w) \quad . \qquad 1.6$$

If such an a exists, the stock should then be raised to the level which will minimize

$$c \cdot a + \alpha\int V(a,w')\,d\varphi(w'/w) \qquad 1.7$$

over $(z,\bar{x}]$.

The optimal policy and value functions for this problem will be examined. The proof of the results will employ K-convex functions, introduced by Scarf [19], p. 199 .

<u>Definition</u> (Scarf) . Let $K \geq 0$. The function f is <u>(strictly)</u> <u>K-convex</u> if

$$K + f(y+x) - f(x) - y\left[\frac{f(x) - f(x-b)}{b}\right] \ (>) \ \geq \ 0 \qquad 1.8$$

where $y > 0$, $b > 0$.

Scarf also states a number of useful properties of K-convex functions:

i) 0 - convexity is equivalent to ordinary convexity.

ii) If $f(x)$ is K - convex, then $f(x+y)$ is K - convex for all y .

iii) If f and h are K - convex and H - convex, respectively, then $\rho f + \sigma h$ is $(\rho K + \sigma H)$ - convex when ρ and σ are positive.

As pointed out by Scarf (iii) extends to integrals when the exchange of limits is permissible. Using these properties it is easy to show the following result (which was presented by Reed [18] for non-differentiable K - concave functions).

<u>Theorem</u> (Reed) Suppose $f(x)$ is continuous and strictly K - convex on $[\underline{x},\bar{x}]$, and let $N = \inf_{[\underline{x},\bar{x}]} f(x)$ and $\bar{S} = \sup\{y : f(y) = N , y \in [\underline{x},\bar{x}]\}$.

Then there exists at most one $\bar{s}$, $\underline{x} \leq \bar{s} \leq \bar{S}$, such that $f(\bar{s}) = N+K$, and further, if such an $\bar{s}$ exists

$$f(x) > N + K \quad \text{for} \quad x \in [\underline{x},\bar{s}] \ .$$

Given this result, the following theorem for the infinite horizon problem may be directly obtained using the contraction operator approach to dynamic programming:

<u>Theorem 1</u> Given assumptions 1 to 3, the following results hold:

i) For $K > 0$, the optimal value function V exists and is continuous and bounded in x and w and is K-convex in x. The optimal policy takes the form:

$$a = \begin{cases} z & \text{if } z > \bar{s}(w) \\ \bar{S}(w) & \text{if } z \leq \bar{s}(w) \end{cases} \tag{1.9}$$

where $z = g(x,w)$. The maps $\bar{s}(\cdot)$, $\bar{S}(\cdot)$ are measurable in w and $\underline{x} \leq \bar{s}(w) \leq \bar{S}(w) \leq \bar{x}$ for all w.

ii) For $K = 0$, the optimal value function V exists and is continuous and bounded in x and w and is convex in x. The optimal policy takes the form

$$a = \begin{cases} z & \text{if } z > \bar{S}(w) \\ \bar{S}(w) & \text{if } z \leq \bar{S}(w) \end{cases} \tag{1.10}$$

where $z = g(x,w)$ and $\underline{x} \leq S(w) \leq \bar{x}$.

For the case of independent and identically distributed disturbances the value function V may be written as

$$V(x,w) = \min_{a \geq z} [h(z) + \delta_{(a-z)} \cdot K + c \cdot (a-z) + \alpha \int V(g(a,w'))\, d\varphi(w')] . \tag{1.11}$$

For the i.i.d. case we immediately obtain the following corollary to Theorem 1 .

Corollary Given assumptions 1 to 3 ,

i) For $K > 0$, the value function $V : X \to \mathbb{R}$ is continuous, bounded, and K - convex. The optimal policy takes the form

$$a = \begin{cases} z & \text{if } z > \bar{s} \\ \bar{S} & \text{if } z \leq \bar{s} \end{cases} \tag{1.12}$$

where $z = g(x,w)$ and $\bar{s}, \bar{S}$ are constants such that $\underline{x} \leq \bar{s} \leq \bar{S} \leq \bar{x}$.

ii) For $K = 0$, the value function $V : X \to \mathbb{R}$ is continuous, bounded, and convex. The optimal policy takes the form

$$a = \begin{cases} z & z > \bar{S} \\ \bar{S} & z \leq \bar{S} \end{cases} \tag{1.13}$$

where $z = g(x,w)$ and $\bar{S}$ is a constant such that $\underline{x} < \bar{S} \leq \bar{x}$.

The proof will be given after three lemmata are stated and proven. Let $B(X \times E)$ be the space of bounded and measurable functions on $X \times E$. The space $B(X \times E)$ is Banach under the sup norm. Let $C_K(X \times E)$ be the space of all bounded, continuous functions on $X \times E$ which are also K - convex on X .

Define the operator $T : B(X \times E) \to B(X \times E)$ by

$$(Tv)(x,w) = \min_{a \in [z,x]} [h(z) + \delta_{(a-z)} \cdot K + c \cdot (a-z) + \alpha \int v(a,w')\, d\varphi(w'/w)]$$

where $v \in B(X \times E)$ and $z = g(x,w)$.

The motivation behind this theorem is that it is easy to see that T is a contraction on the Banach Space $B(X \times E)$ since T is a linear operator there. Then, in any closed subset of the Banach Space which is mapped into itself by T, there must be a fixed point.

lemma 1: The space $C_K(X \times E)$ is complete.

Proof: We need only show that $C_K(X \times E)$ is a closed subset of the complete space $B(X \times E)$. It is sufficient to show that $(f_n) \subset C_K(X \times E)$ and $f_n \to f$ in the sup norm implies that $f \in C_K(X \times E)$.

Clearly f is a continuous bounded function, so we need to show that f is K-convex. Since $f_n \in C_K(X \times E)$

$$K + f_n(x+y) - f_n(x) - y\left[\frac{f_n(x) - f_n(x-b)}{b}\right] \geq 0 \quad \forall\, x,y,b .$$

$f_n \to f$ implies $f_n(x) \to f(x)$, $f_n(x+y) \to f(x+y)$, and $f_n(x-b) \to f(x-b)$ hence

$$K + f(x+y) - f(x) - y\left[\frac{f(x) - f(x-b)}{b}\right] \geq 0 .$$

Thus $f \in C_K(X \times E)$.

lemma 2

For $K > 0$ the operator T on $C_K(X \times E)$ defined by

$$(Tv)(x,w) = \min_{a \in (z,\bar{x}]} [h(z) + \delta_{(a-z)} \cdot K + c \cdot (a-z) + \alpha \int v(a,w')\, d\varphi(w'/w)]$$

where $v \in C_K(X \times E)$ and $z = g(x,w)$, takes the space $C_K(X \times E)$ into itself.

Proof: Clearly (Tv) is bounded in (x,w) . Since v is continuous and K - convex in x then by property (iii) of K - convex functions the term

$$\alpha \int v\,(a,w')\, d\varphi(w'/w)$$

is continuous and strictly K-convex in a . So, the sum

$$[c \cdot a + \alpha \int v(a,w')\, d\varphi(w'/w)]$$

is continuous and strictly K-convex in a . Thus the conditions of Reed's Theorem are satisfied. Let $N(w)$ and $\bar{S}(w)$ be defined by

$$N(w) = \inf_{a \in [\underline{x},\bar{x}]} [c \cdot a + \alpha \int v(a,w')\, d\varphi(w'/w)]$$

and

$$\bar{S}(w) = \sup\{y : [c \cdot y + \alpha\int v(y,w')\, d\varphi(w'/w)] = N(w),\ y \in [\underline{x},\bar{x}]\ \} .$$

(Note that by the maximum principle (see Berge [4]) , $N(w)$ is continuous in w). Then there exists at most one $\bar{s}(w)$, where $\underline{x} \le \bar{s}(w) \le \bar{S}(w) \le \bar{x}$ such that

$$c \cdot \bar{s}(w) + \alpha\int v(\bar{s}(w),w')\, d\varphi(w'/w) = N(w) + K$$

and further, if such an $\bar{s}(w)$ exists,

$$c \cdot y + \alpha\int v(y,w')\, d\varphi(w'/w) > N(w) + K$$

for $y \in [\underline{x},\bar{s}(w))$. If such an $s(w)$ does not exist let $\bar{s}(w) = \underline{x}$. The maps $\bar{s}(w)$, $\bar{S}(w)$ will be measurable in w . Thus (Tv) may be rewritten as

$$(Tv)(x,w) = \begin{cases} h(z) + \alpha\int v(z,w')\, d\varphi(w'/w) & \text{if } z > \bar{s}(w) \\ h(z) + K + c\cdot(\bar{S}(w) - z) + \alpha\int v(\bar{S}(w),w')\, d\varphi(w'/w) & \text{if } z \le \bar{s}(w) \end{cases}$$

where $z = g(x,w)$. Substituting for z ,

$$(Tv)(x,w) = \begin{cases} h(g(x,w)) + \alpha\int v(g(x,w),w')\, d\varphi(w'/w) & \text{if } z > \bar{s}(w) \\ h(g(x,w) + K + c\cdot(\bar{S}(w) - g(x,w)) + \alpha\int v(\bar{S}(w),w')\, d\varphi(w'/w) & \text{if } z \le \bar{s}(w) . \end{cases}$$

(Tv) will now be shown to be continuous in x and w . The term

$$h(g(x,w)) + \alpha\int v(g(x,w),w')\, d\varphi(w'/w)$$

is continuous in x and w by the continuity of $h(\cdot), g(\cdot,\cdot)$, $v(\cdot,\cdot)$

and the continuity of $\varphi(\cdot/w)$ in the weak topology. Also, the term

$$h(g(x,w)) + K - c \cdot g(x,w)$$

will be continuous in x and w. From the definition of $N(\cdot)$

$$c \cdot \bar{S}(w) + \alpha\int v(\bar{S}(w),w')\, d\varphi(w'/w) = N(w)$$

is continuous in w. Since the term is constant in x, it is continuous in x. Thus, the expression

$$h(g(x,w)) + K + c \cdot (\bar{S}(w) - g(x,w)) + \alpha\int v(\bar{S}(w),w')\, d\varphi(w'/w)$$

is continuous in x and w.

So (Tv) is continuous for those x and w where $g(x,w) \leq \bar{s}(w)$ and where $g(x,w) > \bar{s}(w)$. Further, since

$$c\cdot\bar{s}(w) + \alpha\int v(\bar{s}(w),w')\, d\varphi(w'/w) = N(w) + K = c\cdot\bar{S}(w) + \alpha\int v(\bar{S}(w),w')\, d\varphi(w'/w) + K$$

or

$$\alpha\int v(\bar{s}(w),w')\, d\varphi(w'/w) = c\cdot(\bar{S}(w) - \bar{s}(w)) + K + \alpha\int v(\bar{S}(w),w')\, d\varphi(w'/w)$$

then at $g(x,w) = \bar{s}(w)$, (Tv) is continuous in x,w. So (Tv) is continuous in x,w.

Let w be fixed. For x such that $g(x,w) > \bar{s}(w)$, the function (Tv) is composed of the term

$$h(g(x,w))$$

which is convex for all $x \in [\underline{x},\overline{x}]$ since $h(\cdot)$ is convex and g is linear in x. Also $\alpha\int v(g(x,w),w')\,d\varphi(w'/w)$ is strictly K-convex since v is K-convex and $0 < \alpha < 1$. For x such that $g(x,w) \leq \overline{s}(w)$, (Tv) is composed of a constant term and the convex function

$$h(g(x,w)) - c \cdot g(x,w)$$

so (Tv) is K-convex for the intervals on $[\underline{x},\overline{x}]$ where $g(x,w) \geq \overline{s}(w)$ and $g(x,w) < \overline{s}(w)$. Consider the interval $[x,y]$ where x is such that $g(x,w) \leq s(w)$ and $g(y,w) > s(w)$. Let $y = x + a$. We wish to show that for w chosen arbitrarily:

$$K + (Tv)(a + x,w) - (Tv)(x,w) - a\left[\frac{(Tv)(x,w) - (Tv)(x - b,w)}{b}\right] \geq 0$$

where $a,b > 0$.

$$K + (Tv)(a + x,w) - (Tv)(x,w) - a\left[\frac{(Tv)(x,w) - (Tv)(x - b,w)}{b}\right]$$

$$= K + [h(g(a+x,w)) + \alpha\int v(g(a+x,w),w')\,d\varphi(w'/w)\ ]$$

$$- [h(g(x,w)) + K + c \cdot (S(w) - g(x,w)) + \alpha\int v(S(w),w')\,d\varphi(w'/w)]$$

$$- \frac{a}{b}\ [h(g(x,w)) + K + c \cdot (S(w) - g(x,w)) + \alpha\int v(S(w),w')\,d\varphi(w'/w)$$

$$- h(g(x-b,w)) + K + c(S(w) - g(x-b,w)) + \alpha\int v(S(w),w')\, d\varphi(w'/w) \quad]$$

$$= \left\{ h(g(a+x,w)) - h(g(x,w)) - a \left[\frac{h(g(x,w)) - h(g(x-b,w))}{b} \right] \right\}$$

$$+ \left\{ [c \cdot g(a+x,w) + \alpha\int v(g(a+x,w),w')\, d\varphi(w'/w)] - [c \cdot S(w) + \alpha\int v(S(w),w')\, d\varphi(w'/w)] \right\}$$

$$+ c \cdot \left\{ g(x,w) - g(a+x,w) + \frac{a}{b}[g(x,w) - g(x-b)] \right\} .$$

The first term is non-negative since $h \circ g$ is convex in x. The second term is non-negative by the definition of $S(w)$. Since g is linear, the third term will be zero. So (Tv) will be K - convex for all $x \in [\underline{x},\overline{x}]$. This will be true for any w since w was chosen arbitrarily.

So $T : C_K(X \times E) \rightarrow C_K(X \times E)$.

<u>lemma 3</u> Let $T : B(X \times E) \rightarrow B(X \times E)$ be defined as in the text by

$$(Tv)(x,w) = \min_{a[z,\overline{x}]} [h(z) + \delta_{(a-z)} \cdot K + c(a-z) + \alpha\int v(a,w')\, d\varphi(w'/w] .$$

Then there exists a $V^* \in C_K(X \times E)$ such that $TV^* = V^*$.

Proof: Since T is clearly a monotone, linear operator and $0 < \alpha < 1$ it is a contraction mapping on $B(X \times E)$. Since $T : C_K(X \times E) \rightarrow C_K(X \times E)$ and $C_K(X \times E) \subset B(X \times E)$, T is also a contraction on $C_K(X \times E)$. Hence by the Contraction Mapping Principle there is a unique point $V^* \in C_K(X \times E)$ such that $TV^* = V^*$, see Liusternik and Sobolev [16].

proof of the theorem:

i) The argument now follows the standard dynamic programming framework (see for example Maitra [17] and Hinderer [10]). With each Borel measurable map

$$a : X \times E \to X$$

associate the operator $Q(a) : C_K(X \times E) \to C_K(X \times E)$ defined by

$$(Q(a)v)(x,w) = h(z) + \delta_{(a(x,w) - z)} \cdot K + c \cdot (a(x,w) - z) + \alpha \int v(a(x,w)\, d\varphi(w'/w) \quad \text{for}$$

for $K > 0$.

Since $Q(a)$ is a monotone, linear operator and $0 < \alpha < 1$ it is a contraction on $C(X \times E)$ (the space of continuous and bounded functions on $X \times E$) and thus has a unique fixed point in $C_K(X \times E)$ by an argument similar to that for T. Let $I(a)$ be this unique fixed point. By the Selection Theorem of Dubins and Savage [8], (see Maitra [17]), there is a Borel measurable map $a : X \times E \to X$ such that $TV = Q(a)V$. Hence $Q(a)V = V$ and $V = I(a)$ since $I(a)$ is the unique fixed point of $Q(a)$. Therefore $TV = V$ can be rewritten as

$$I(a)(x,w) = \min_{a \in [z,\overline{x}]} [h(z) + \delta_{(a-z)} \cdot K + c \cdot (a - z) + \alpha \int v(a,w')\, d\varphi(w'/w)]$$

for $z = g(x,w)$. Thus $I(a)$ satisfies the optimality equation and the stationary plan $(a,a,a,\ldots)$ is an optimal plan (see Blackwell [6]).

As $V = I(a)$ and $V \in C_K(X \times E)$ the optimal return is continuous and bounded on $X \times E$ and K-convex on X. Since $0 < \alpha < 1$, the term that the planner seeks to minimize, if the decision is made to restock,

$$c \cdot a + \alpha \int V(a,w')\, d\varphi\, (w'/w)$$

is continuous and strictly K - convex in a by properties (i) and (iii) of K-convex functions. Hence the conditions of Reed's theorem are satisfied and (by arguments similar to those used in the appendix) the optimal policy will have the form

$$a = \begin{cases} z & \text{if } z > \bar{s}(w) \\ \bar{S}(w) & \text{if } z \leq \bar{s}(w) \end{cases}.$$

ii) When $K = 0$, the contraction operator T may be defined on $C(X \times E)$ the space of continuous, bounded functions on $X \times E$ which are convex on X. T is defined by

$$(Tv)(x,w) = \min_{a \in [g(x,w),\bar{x}]} [h(g(x,w)) + c \cdot (a - g(x,w)) + \alpha \int v(a,w')\, d\varphi(w'/w)] .$$

Proceeding as before, the optimal value function V can be shown to be continuous and bounded on $X \times E$ and convex on X. Also, an optimal plan

$$I(a)(x,w) = \min_{a \in [g(x,w),\bar{x}]} [h(g(x,w)) + c \cdot (a - g(x,w)) + \alpha \int V(a,w')\, d\varphi(w'/w)]$$

where the term

$$c \cdot a + \alpha \int V(a,w')\, d\varphi(w'/w)$$

is continuous and convex in a.

$$\text{Let } N(w) = \inf_{a \in [\underline{x},\overline{x}]} [c \cdot a + \alpha \int V(a,w')\, d\varphi(w'/w)]$$

$$\text{and } S(w) = \sup \{y : [c \cdot y + \alpha \int V(y,w')\, d\varphi(w'/w)] = N(w), \quad y \in [\overline{x},\overline{x}]\} .$$

These are well defined since the bracketed term is continuous and convex. (Note that $N(\cdot)$ is continuous by the Berge Maximum Principle, see [4]). Then

$$c \cdot S(w) + \alpha \int V(S(w),w')\, d\varphi(w'/w) = N(w)$$

so that

$$c \cdot S(w) + \alpha \int V(S(w),w')\, d\varphi(w'/w) \leq c \cdot a + \alpha \int V(a,w')\, d\varphi(w'/w) .$$

Thus the optimal policy has the form

$$a = \begin{cases} z & \text{if } z > \overline{S}(w) \\ \overline{S}(w) & \text{if } z \leq \overline{S}(w) \end{cases} .$$

Note that $\overline{S}(w)$ will be measurable .

II. Regeneration Theorem

In this section we provide a theorem, similar to the theorems of renewal theory, about the convergence of a particular type of Markov Process. The type of process that we consider has a regeneration point--a point which the process attains with strictly positive probability, regardless of its initial state. Such processes naturally result from the interaction of exogenous randomness and optimal policies in many decision problems. It will be seen in the next section that the Markov Processes resulting from the inventory problem have this form in most cases. The theorem will allow consideration of the asymptotic properties of the process of inventory levels without explicit consideration of a renewal equation (see [14] and [22] for a discussion of the Renewal Theory approach).

The state space of the system will be denoted by the measure space $(S,\mathscr{S})$ where S is a complete separable metric space and $\mathscr{S}$ is its Borel σ-algebra. The transition probability on $S \times \mathscr{S}$ will be denoted by $P : S \times \mathscr{S} \to [0,1]$. $P(s,\cdot)$ is the probability measure on tomorrow's states when the current state is $s \in S$ and is derived from an individual decision problem. The following definitions will be used frequently in this section:

1) $ca(S)$ -- the set of all bounded, real valued, countably additive set functions with domain $\mathscr{S}$.

2) $P(S)$ -- the subset of $ca(S)$ consisting of the probabilities on S.

Note: all spaces of set functions will be endowed with the weak topology. That is, for $(\mu_n) \subset ca(S)$ $\mu_n \to \bar{\mu}$ if and only if

$$\int f(s)\mu_n(ds) \to \int f(s)\bar{\mu}(ds)$$

for all real valued, bounded, continuous functions f on S.

The transition probability P defines a linear operator M^* mapping $ca(S)$ into itself, by

$$M^*\lambda(A) = \int P(s,A)\lambda(ds) \qquad \begin{array}{l} \text{for } A \in \mathscr{S} \\ \text{for } \lambda \in ca(S). \end{array}$$

M^{*n} is defined in the same way from P^n the n-step transition probability. Then $M^{*n}\lambda$ gives the probability measure that will result in n periods of time if the initial probability is λ.

The following two definitions will be used to state the assumptions needed in the following convergence theorem.

Definition: A point $\rho \in S$ will be called a regeneration point if there is an integer $n > 0$ and a $\beta > 0$ such that $P^n(s,A) \geq \beta > 0$ for all $s \in S$ and for all measurable sets A containing ρ.

That is, if ρ is a regeneration point, then there is a strictly positive probability of reaching ρ in a finite number of steps from any point in the state space. The process will then start over from ρ since it is Markov.

Definition: A Markov Process will be called stochastically bounded if there is an integer n such that the family of probability measures $\{M^{*n}\lambda, \lambda \in P(S)\}$ is tight, i.e. if for any $\epsilon > 0$ there is a compact set $C \subseteq S$ such that $M^{*n}\lambda(c) \geq 1 - \epsilon$ for all $\lambda \in P(S)$.

That is, there must be a compact set such that eventually the process stays in this set most of the time. Clearly any process on a compact state space is stochastically bounded.

Theorem 2 Any stochastically bounded Markov Process with a regeneration point has an invariant probability λ^* and $M^{*n}\lambda$ converges to λ^* for any initial probability λ .

Proof: We show that this system satisfies Doeblin's Condition and has only one compact ergodic set with no cyclically moving subsets. These results will imply the conclusions of the theorem, see Doob [7] .

We will first show that the linear operator M associated with P is quasi-weakly compact. An operator is quasi-weakly compact if and only if the corresponding transition probability satisfies Doeblin's Condition, see Futia [9] , Theorems 4.4 and 4.9. A linear operator $M : ca(S) \to ca(S)$ is quasi-weakly compact if there is a linear operator $L : ca(S) \to ca(S)$ with closure $(L(P(S)))$ weakly compact and $\|M^{*n} - L\| < 1$ for some integer n .

Let $L : ca(S) \to ca(S)$ be defined by

$$L\lambda(A) = \begin{cases} \beta|\lambda| & \text{if } \rho \in A \\ 0 & \text{otherwise.} \end{cases}$$

L is clearly a linear operator. We want to show that closure $(L(P(S)))$ is weakly compact. Note that $L(P(S)))$ can be identified with the compact interval $[0,\beta]$ by $f : [0,\beta] \to L(P(S))$ defined by

$$f(c) = \mu \quad \text{such that} \quad \mu(A) = \begin{cases} c & \rho \in A \\ 0 & \rho \notin A \end{cases} .$$

Hence $L(P(S))$ is the image of the compact set $[0,\beta]$ under a mapping f which is clearly weakly continuous. Therefore $L(P(S))$ is weakly compact.

We now want to show that $\| M^{*n} - L \| < 1$.

$$\| M^{*n} - L \| = \sup_{\lambda \in P(S)} |M^{*n}\lambda - L\lambda| = \sup_{\lambda \in P(S)} |M^{*n}\lambda - \beta\delta_\rho| \leq 1 - \beta < 1$$

since $M^{*n}\lambda(A) \geq \beta$ for all A such that $\rho \in A$ and $M^{*n}\lambda(A) \leq 1 - k$ for all other $A \in \mathcal{S}$. Therefore Doeblin's Condition is satisfied.

Ergodic sets must be essentially disjoint and for any ergodic set E, $P(s,E) = 1$ for all $s \in E$. Since $P^n(s,A) \geq \beta > 0$ for all A containing ρ and for all $s \in S$ there can be at most one ergodic set. There is at least one since the process is stochastically bounded.

A sequence of cyclically moving sets $C_1, C_2, \ldots, C_m$ must be essentially disjoint and $P(s, C_{i+1}) = 1$ for any $s \in C_i$, for all i. Suppose there is such a sequence of sets. Consider $s \in C_i$, $P^n(s,A) \geq \beta > 0$ for all A containing ρ implies that $\rho \in C_{i+n}(\text{mod } m)$. Repeating for all i we have $\rho \in C_i$ for all i and hence the sets C_i are not essentially disjoint. Therefore there are no cyclically moving sets.

In the following section we show that under certain assumptions, in both the i.i.d. and Markov cases, the transition equation resulting from an optimal inventory policy generates a transition probability which satisfies the assumptions of this theorem and hence has an invariant probability reflecting the steady state behavior of inventories. The unifying element of such problems is that the optimal policy has a very particular structure resulting from linearity of the function representing current costs and benefits from altering the stock. The presence of non-convexities in the objective function will not affect the solution since the optimal policy will be to do nothing over some segment of the state space and then at the boundary of this segment jump to a particular state. This state will be shown to be a regeneration state. Such a regeneration state will result from any policy, whether optimal or not, which has this property.

III. Convergence to a Steady State Distribution

The asymptotic properties of the inventory process subject to Markov disturbances will now be examined. Repeated application of the optimal stock adjustment policy $a: X \times E \to X$ in the presence of random disturbances will generate a stochastic process on the space of inventories X. We may define a stochastic process on the state of the system $S = X \times E$. The transition probability measure

$$P : S \times \mathscr{S} \to [0,1] \tag{3.1}$$

will be given by

$$P(s,B) = \delta_{a(s)}(x') \circ \varphi\,(\operatorname*{proj}_E B/w)$$

for $s \in S$ and $B \in \mathscr{S}$, for the Markov case. For the i.i.d case, the state space can be restricted to X and the transition probability can be written as

$$P(x,B) = \varphi(w \in E/a(x,w) \in B) \tag{3.2}$$

for $x \in X$ and $B \in X$. P gives the probability measure on the next state of the system that results from the composition of exogenous randomness with the optimal plan given the current state of the system. It is easy to show that P defines a Markov process on $(S,\mathscr{S})$.

The results of Sections I and II will be used to show that the sequence of probability measures over time converges to a unique probability regardless of the initial state. This invariant probability can be interpreted as the result of observing one decision maker over a long period of time or of observing a cross section of independent identical decision makers at one time. We will use Theorem 2 and show also that the optimal inventory $\bar{S}(\bar{w})$ forms a regeneration point. Before proceeding however, the following additional condition is placed jointly on the transition equation g and measure φ to assure an interesting problem:

<u>Condition 1</u> a) When disturbances are Markov there exists a disturbance $\bar{w} \in E$ satisfying

$$\varphi(\bar{w} \mid w) \geq \gamma > 0$$

for all $w \in E$, such that for a sequence (z_n) defined by $z_1 = g(x_o,\bar{w}), z_2 = g(z_1,\bar{w}), z_3 = g(z_2,\bar{w})$, for all x_o, there exists an integer N such that $z_N \leq \underline{x}$.

b) When disturbances are i.i.d., there exists a set $A \subset E$ satisfying

$$\varphi(A) \geq \gamma > 0$$

such that for the sequence (z_n) defined by $z_1 = g(x_o,w_o), z_2 = g(z_1,w_1)$, $z_3 = g(z_2,w_2),\ldots,$ for all x_o, and for any $w_n \in A$, there exists an integer N such that $z_N \leq \underline{x}$.

The effect of this condition may be examined for the inventory problem. The restriction for the i.i.d. case simply implies that when the stock is not replenished, demand will exhaust the current stock to zero (or to the limit on backorders) within a finite number of periods. For the Markov case, the restriction guarantees that at least one demand level will reoccur with positive probability such that the stock will be exhausted within a finite number of periods. These restrictions simply require the presence of a decision maker to periodically readjust the stock and are useful in examining the existence of a regeneration state.

Given Condition 1, it is now straightforward to demonstrate convergence to a unique steady state distribution on the process of inventory levels.

Theorem 3 The sequence of probability measures on inventories converges to a unique probability measure λ^* with $\overline{S}(\overline{w})$ as a regeneration point in the Markov case and $\overline{S}$ as a regeneration point in the i.i.d. case.

Proof: By Theorem 2 we need only show that the process is stochastically bounded and that it has a regeneration point. The process is clearly bounded since S is compact.

We now show that in the Markov case $(\overline{S}(\overline{w}),\overline{w})$ is a regeneration point. By Condition 1(a) there is an $\overline{w} \in E$ with $\varphi(\overline{w}/w) \geq \gamma > 0$ and an N such that $Z_N \leq X$, for any X_0, if no action is taken. Hence from any initial point $(X_0,w) \in S$ the process will reach

$(\underline{x},\overline{w})$ with probability γN if no action is taken. But by Theorem 2 we know that the optimal policy is to restock to $\overline{S}(\overline{w})$ once $\overline{s}(\overline{w}) \geq \underline{x}$ (or $\overline{S}(\overline{w})$ in the case of $K = 0$) is reached. Hence with probability greater than or equal to $\gamma(N + 1)$ the point $(\overline{S}(\overline{w}),\overline{w})$ is reached in at most $N + 1$ steps from any initial point. So $(\overline{S}(\overline{w}),\overline{w})$ is a regeneration point. Therefore there is an invariant probability η on S . The invariant probability λ^* on inventories is then defined by

$$\lambda^*(A) = \eta(a^{-1}(A)) \quad \text{for all} \quad A \in X .$$

Similarly in the i.i.d. case $\overline{S}$ is a regeneration point (by condition 1(b)) in the restricted state space S and hence there is an invariant probability λ^* on s .

Hence any stock management problem fitting the assumptions of this section will lead to a regeneration point and an equilibrium probability on the stocks. For example, in the traditional inventory problem with transition equation $z_t = x_t - w_t$ there is an equilibrium distribution of inventories λ^* on $[\underline{x},\overline{x}]$ in which the optimal inventory level $\overline{S}$ (or $\overline{S}(\overline{w})$) has probability of at least $\gamma(N + 1)$. This equilibrium distribution of inventories could also be easily transformed to give the distribution of costs for the firm.

There will also be an equilibrium distribution of savings in the portfolio problem with transition equation $z_t = w_t^1 x_t + w_t^2$. This distribution describes the observed distribution of savings or money holding among individuals as a result of random needs and returns and an attempt to minimize non-convex costs.

This method can also be applied to the problem of capital management for a firm or investment for an economy in which there are both marginal and fixed or adjustment costs. In this case a possible transition equation is $z_t = f(w_t)x_t$ where x_t is capital stock (or capital labor ratio) and $f(w_t)$ represents random effects of usage rates and depreciation (or random population growth and capital depreciation). In either case there will be an equilibrium distribution of stocks as a result of the planner's actions. Hence in the case of a firm managing a capital stock otherwise identical profit maximizing firms need not have the same capital stocks, although they will have the same distribution of capital over time.

BIBLIOGRAPHY

1. Arrow, K.J., T. Harris, and J. Marshak: "Optimal Inventory Policy," Econometrica, XIX (1951), 250-272.

2. Arrow, K.J., S. Karlin and H. Scarf: Studies in the Mathematical Theory of Inventory and Production, Stanford University Press, Stanford, California, 1958.

3. Arrow, K.J., S. Karlin, and P. Suppes: Mathematical Methods in the Social Sciences, Stanford University Press, Stanford, California, 1960.

4. Berge, C.: Topoligical Spaces, Macmillan, New York, 1963.

5. Bertsekas, Dimitri P.: Dynamic Programming and Stochastic Control, Academic Press, New York 1976.

6. Blackwell, D.: "Discounted Dynamic Programming," Annals of Mathematical Statistics, (36) 1965, 226-235.

7. Doob, J.L.: Stochastic Processes, Wiley and Sons, New York, 1953.

8. Dubins, L.E., and L.J. Savage: How to Gamble if You Must, McGraw-Hill, New York, 1965.

9. Futia, C.: "A Stochastic Approach to Economic Dynamics," Bell Laboratories, New York , (1976).

10. Hinderer, K.: Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter, Lecture Notes in Operations Research and Mathematical Systems, 33 (1970), Springer-Verlag.

11. Hordijk, A. and Tijms, H.: "Convergence Results and Approximations for Optimal (s,S) Policies," Management Science, 20 (1974), 1432-1438.

12. Iglehart, D.L.: "Optimality of (s,S) Policies in the Infinite Horizon Dynamic Inventory Problem," Management Science 9 (1963), 259-267.

13. Karlin, S.: "Steady State Solutions," Studies in the Mathematical Theory of Inventory and Production, Stanford University Press, Stanford, California, 1958, 223-264.

14. _________: "The Application of Renewal Theory to the Study of Inventory Policies," Studies in the Mathematical Theory of Inventory and Production, Stanford University Press, Stanford, California, 1958, 270-297.

15. _________, and A. J. Fabens: "A Stationary Inventory Model with Markovian Demand," Mathematical Methods in the Social Sciences, Stanford University Press, Stanford, California 1960, 159-175.

16. Liusternik, L.A., and Sobolev, V.J.: Elements of Functional Analysis, Wiley, New York, 1974.

17. Maitra, A.: "Discounted Dynamic Programming on Compact Metric Spaces," SANKHYA, Series A, 30 Part 2, (1968), 211-216.

18. Reed, W.J.: "A Stochastic Model of the Economic Management of a Renewable Animal Resource", Mathematical Biosciences 22, (1974)

19. Scarf, H.: "The Optimality of (S,s) Policies in the Dynamic Inventory Problem," in Mathematical Methods in the Social Sciences, Stanford University, Stanford, California, 1960.

20. Scarf, H.E., Gilford, D.M. and Shelly, M.W.: Multistage Inventory Models and Techniques, Stanford University Press, Stanford, 1963.

21. Strauch, R.: "Negative Dynamic Programming," Annals of Mathematical Statistics, 37 (1966), 871-890.

22. Veinott Jr., A.F. and Wagner, A.M.: "Computing Optimal (s,S) Inventory Policies," Management Science, 11 (1965), 525-552.