MATHEMATICS OF OPERATIONS RESEARCH
Vol. 11, No. 2, May 1986
Printed in U.S.A.

AN INVENTORY MODEL WITH LIMITED PRODUCTION CAPACITY AND UNCERTAIN DEMANDS I. THE AVERAGE-COST CRITERION*

A. FEDERGRUEN AND P. ZIPKIN

Columbia University

This paper considers a single-item, periodic-review inventory model with uncertain demands. In contrast to prior treatments of this problem we assume a finite production capacity per period. Assuming stationary data, a convex one-period cost function and a discrete demand distribution, we show (under a few additional unrestrictive assumptions) that a modified base-stock policy is optimal under the average-cost criterion; in addition, we characterize the optimal base-stock level.

1. Introduction. This paper treats one of the most basic of production (or inventory) models, in which the stock of a single item must be controlled under periodic review: The production cost is linear in the amount produced at the beginning of each period, and there is a convex function representing expected penalty and holding costs at the end of the period. Demand in each period is nonnegative and independent of those in other periods. All data are stationary, and the planning horizon is infinite. All stockouts are backordered. Here we assume demands are discrete (integer-valued), and the long-run-average expected-cost criterion. (A companion paper [10] treats continuous demands and discounted costs.)

Prior treatments of this problem assume unlimited production capacity in each period (cf., e.g., Heyman and Sobel [16], Scarf [22], Iglehart [17], Veinott [25]). In this case a stationary base-stock policy, described by a single critical number, is optimal: When initial stock is below that number, enough should be produced to bring total stock up to the number; otherwise, nothing should be produced. This assumption of limitless capacity, while a reasonable approximation in some situations, is a poor one in others.

Here we assume a finite production capacity. In this setting it is plausible that a modified base-stock policy would work well: Follow a base-stock policy when possible; when the prescribed production quantity would exceed the capacity, produce to capacity. We shall call this a base-stock, or critical-number, policy for short. Our goal here is to prove that such a policy is optimal under certain reasonable assumptions. (Compared to the uncapacitated case, the computation of the average cost of such a policy and the determination of an optimal one appear to be difficult; we defer a detailed discussion to a subsequent work.)

If the production costs have a fixed (as well as a variable) component it might be reasonable to expect that a modified (s, S)-policy is optimal: When the inventory level falls below a critical number s, produce enough to bring total stock up to S or as close to it as possible, given the limited capacity; otherwise do not produce. In the uncapacitated case this structure is optimal both for finite- and infinite-horizon problems, cf. Scarf [22] and Iglehart [17]. However, Wijngaard [26] gives an example of

*Received December 8, 1983; revised February 20, 1985.
AMS 1980 subject classification. Primary: 90B05.
IAOR 1973 subject classification. Main: Inventory.
OR/MS Index 1978 subject classification. Primary: 346 Inventory/production/policies.
Key words. Inventory model, periodic review, production policy, base-stock policy.

0364-765X/86/1102/0193$01.25

a finite-horizon model having a more complex optimal policy. Whether the modified (s, S) structure prevails in the infinite-horizon case remains an open question. The analysis in this paper can, however, straightforwardly be used to conclude that, even with fixed plus variable order costs, solutions to the optimality equation exist, and that any stationary policy satisfying the equation for certain such solutions is optimal.

Karlin [18] treats the case of unlimited capacity but a nonlinear production cost, and the current problem can be approximated by one having a production cost function which rises steeply for values greater than the capacity. However, [18] explicitly assumes a finite (indeed differentiable) production cost, and the fundamental qualitative nature of the problem seems to change when a constraint on production is imposed. Wijngaard [26], [27] deals with the capacitated model and addresses the optimality of a (modified) base-stock policy within a restricted class of stationary policies and under different assumptions from ours regarding the one-period cost function and the demand distribution.

The management of an inventory under finite capacity and the mathematical analysis of the problem are complicated by the same phenomenon: In the uncapacitated case the effect of a large demand in some period can be corrected immediately in the next period. Here, several periods of full production may be required, during which further large demands might occur, requiring still more time to return to a normal stock level, etc. Indeed, it is not obvious that the system is stable enough under *any* policy to have a finite average cost. Our assumptions are essentially sufficient conditions for such stability. (Note that the lost-sales case is conceptually simpler in this sense. This is the problem treated by Evans [7], assuming a finite horizon and multiple items.)

The possible buildup of backorders just described is reminiscent of the behavior of a queue. Indeed, using a standard transformation, the current problem can be shown to be equivalent to a certain continuous-time queueing-control model. (Presuming a finite inventory capacity, as we do sometimes below, the length of the queue is identified with inventory capacity minus inventory.) Customers can be served in batches of size up to the production capacity, with service time independent of batch size. The size of the next batch is decided whenever a service is complete; even if zero customers are served, a "service" commences and must finish before the next batch can begin. This factor distinguishes the model from the work of Deb and Serfozo [3] on control of a batch-service queue under more standard assumptions; also, their cost structure differs from ours. (The optimal policies in the two cases, however, are of the same form.)

There has been considerable research recently on continuous-time production models in which production occurs continuously at a finite rate (cf., e.g., Doshi et al. [5], Gavish and Graves [14], Graves and Keilson [15] and De Kok et al. [4]), or on a unit basis (cf., e.g., Tijms [24]). The problems addressed by these models are clearly closely related to ours, but our assumption of batch production seems to preclude a reduction of one to the other.

We remark that capacity limits in deterministic production-scheduling models have drawn the attention of researchers lately (e.g., Florian and Klein [12], Florian, Lenstra and Rinnooy Kan [13], Bitran and Yanasse [2], Baker et al. [1]). Here too, finite capacity results in complications, the degree of which depends on the details of the problem.

§2 defines notation and presents the required assumptions, and §3 proves some preliminary results used subsequently. In §4 we prove the optimality of a critical-number policy, assuming a finite inventory capacity; the key observation is that, while the finite production capacity complicates the problem in some ways, the set of feasible actions in each state is compact, which permits us to invoke results from Federgruen, Schweitzer and Tijms [9]. In §5 we remove the inventory-capacity assumption and prove that a critical-number policy is optimal within a slightly restricted class of policies, including all stationary policies.

2. Notation and assumptions. D = generic random variable representing one-period demand.

D_t = demand in period t, $t = 0, 1, \ldots$

We assume that D_t are independent, and each has the same distribution as D.

$D^{(i)}$ = generic i-period demand, $i = 1, 2, \ldots$

$p(j) = \Pr\{D = j\}, j = 0, 1, \ldots$

$P(j) = \Pr\{D \leq j\}, j = 0, 1, \ldots$

b = production capacity, or limit on order size, a positive integer.

c = per-unit order (production) cost.

x_t = inventory at the beginning of period t, $t = 0, 1, \ldots$

y_t = inventory after ordering (production) but before demand in period t, $t = 0, 1, \ldots$

Since the problem data are stationary, we suppress the time index where possible, writing x for x_t for example. Both x and y are always integer-valued. Since stockouts are backordered, x and y may be negative.

U = inventory capacity, $U > 0$.

$Y(x) = \{y : x \leq y \leq x + b, y \leq U\}$, the feasible values of y given $x \leq U$.

$G(y)$ = one-period expected cost function (including, for example, expected penalty and holding costs).

For much of the paper U is assumed to be a finite integer. The case $U = +\infty$ is discussed in §5. Nevertheless, G is presumed to be defined for all integers y.

We now state additional assumptions:

ASSUMPTION 1. (a) $\lim_{|y|\to\infty} G(y) = \infty$; (b) *$G$ is nonnegative and convex.*

Then we can define

$$\bar{y}^{\infty} = \text{smallest } y \text{ achieving the global minimum of } G.$$

For convenience we assume $U \geq \bar{y}^{\infty}$, though the results are unaffected otherwise.

ASSUMPTION 2. *The characteristic function of D is analytic in a neighborhood of the origin.*

This implies that D has finite moments of all orders; in particular, let

$$\mu = E(D), \qquad 0 < \mu < \infty.$$

$$\sigma^2 = \operatorname{Var}(D), \qquad 0 < \sigma^2 < \infty.$$

Most of the "standard" demand densities satisfy this assumption. (Among continuous densities, however, the lognormal is a notorious violator, cf. Feller [11, p. 227].)

We add a slightly stronger condition to Assumption 1(a), to preclude it being optimal never to order: $\lim_{y\to-\infty}[cy + G(y)] = \infty$. Then every policy with finite average total cost has the same average order cost $c\mu$, so we set $c = 0$ without loss of generality. (Recall, $\bar{y}^{\infty}$ would be the optimal critical number with unlimited production capacity, cf., e.g., Heyman and Sobel [16].)

ASSUMPTION 3. *For some positive integer ρ, $G(y) = O(|y|^{\rho})$.*

(That is, there are positive constants A and B with $G(y) \leq A + B|y|^{\rho}$ for all y.) Often G is derived as the expectation of linear penalty and holding cost functions. In this case Assumption 3 is satisfied with $\rho = 1$.

ASSUMPTION 4. (a) $b > \mu$; (b) $P(b) < 1$.

In the absence of (a), $E(y_t) \to -\infty$, and the cost is infinite under every policy. Without (b) the problem is essentially uncapacitated.

Denote

$\Phi(\cdot)$ = standard normal cumulative distribution function.

$\phi(\cdot)$ = standard normal density.

Also, let R_E be the real line with the Euclidean topology, and $R_E^U = \times_{x \leqslant U}(R_E)$, a cartesian product endowed with the corresponding product topology (Royden [21, p. 150]). We shall use R_E^U as the set of real-valued functions $v(x)$ defined on the state space, with this useful topological structure. Let $v0 \in R_E^U$ denote the zero-function, $v0(x) = 0$, $x \leqslant U$, and $v1(x) = 1$, $x \leqslant U$. Also, for any positive integer q, let $V_q = \{v : v \in R_E^U, v(x) = O(|x|^q)\}$, and $V_\infty = \bigcup_q V_q$.

Let Δ denote the set of pure, stationary policies for the problem; for each $\delta \in \Delta$, $y = \delta(x)$ denotes the action prescribed by δ in state x. We shall also use $\delta \in \Delta$ to denote a one-period policy, and sequences of δ's to denote memoryless, nonstationary (that is, Markov) policies; the meaning will be clear from the context.

We consider in addition a class of models with restricted actions sets: For each integer $L < U - b$ let

$$Y_L(x) = \begin{cases} \{x + b\}, & x \leqslant L, \\ Y(x), & L < x \leqslant U. \end{cases}$$

Let Δ_L be the set of policies in Δ feasible with respect to these action sets. (One is thus forced to order to capacity when inventory drops to L or below.) Let R denote the operator $Rv(y) = G(y) + Ev(y - D)$, $y \leqslant U$, and define the value-iteration operators for the original and restricted models, respectively, on V_∞:

$$Sv(x) = \inf\{Rv(y) : y \in Y(x)\}, \qquad S_L v(x) = \inf\{Rv(y) : y \in Y_L(x)\}, \qquad x \leqslant U.$$

(The expectations are well defined and finite by Assumption 2. Observe, the optimality equation for the unrestricted model can be written $gv1 + v = Sv$. This equation plays a key role in §4 below.) Also, define the reduced value-iteration operators

$$Qv(x) = Sv(x) - Sv(\bar{y}^\infty), \qquad Q_L v(x) = S_L v(x) - S_L v(\bar{y}^\infty), \qquad x \leqslant U.$$

For any operator A, denote $A^1 = A$ and $A^{n+1}v = A(A^n v)$, $n \geqslant 1$.

Let $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the integer round-down and round-up of $(\cdot)$, respectively.

3. **Preliminary results.** We first state a result of Feller [11, p. 552], used in the sequel. (It is this result which requires Assumption 2.) Let $\{z_t\}_{t=1}^\infty$ be a sequence of real numbers with each $z_t > 0$, $\lim_{t\to\infty}\{z_t\} = \infty$, but $z_t = o(t^{1/6})$.

LEMMA 1 (Feller [11]): *There exists a constant $C > 0$ such that, for all $t \geqslant 1$,*

$$\Pr\left\{\frac{D^{(t)} - t\mu}{\sigma\sqrt{t}} > z_t\right\} \leqslant \left[1 + C\left(z_t^3/\sqrt{t}\right)\right]\left[1 - \Phi(z_t)\right]$$

The following facts will be useful below:

LEMMA 2. (a) $1 - \Phi(z) < \exp(-\frac{1}{2}z)$, $z \geqslant 0$.

(b) *Let a, r be nonnegative numbers and k an integer, $k > a$.*

$$\sum_{t=k+1}^{\infty} \frac{t^2 + rt}{(t-a)^4} \leqslant \frac{1}{k-a} + \frac{r + 2a}{2(k-a)^2} + \frac{ra + a^2}{3(k-a)^3}.$$

(c) *For real numbers* r, β, a, *with* $r \geqslant 0$, $a, \beta > 0$, *let*

$$n_j(t) = j^r \exp(-j^\beta / a\sqrt{t}), \qquad j = 0, 1, \ldots; \quad t = 1, 2, \ldots$$

and $\lambda = (r+1)/\beta$. *Then*, $\sum_{j=0}^{\infty} n_j(t) = O(t^{\lambda/2})$.

PROOF. (a) For $0 \leqslant z \leqslant 1$

$$1 - \Phi(z) \leqslant 1 - \Phi(0) = \tfrac{1}{2} < \exp(-\tfrac{1}{2}) \leqslant \exp(-\tfrac{1}{2}z).$$

For $z > 1$ it is easy to show $1 - \Phi(z) - \phi(z) < 0$ (the function goes to 0 as $z \to \infty$, and its derivative is > 0 for $z > 1$), so $1 - \Phi(z) < \phi(z) < \exp(-\tfrac{1}{2}z)$.

(b) We have

$$\frac{t^2 + rt}{(t-a)^4} = \frac{1}{(t-a)^2} + \frac{r+2a}{(t-a)^3} + \frac{ra + a^2}{(t-a)^4},$$

$$\sum_{t=k+1}^{\infty} (t-a)^{-2} \leqslant \sum_{t=k+1}^{\infty} \int_{t-1}^{t} (z-a)^{-2}\, dz = \int_k^{\infty} (z-a)^{-2}\, dz = (k-a)^{-1},$$

and similarly,

$$(r+2a) \sum_{t=k+1}^{\infty} (t-a)^{-3} \leqslant (r+2a) \int_k^{\infty} (z-a)^{-3}\, dz = \tfrac{1}{2}(r+2a)(k-a)^{-2},$$

$$(ra+a^2) \sum_{t=k+1}^{\infty} (t-a)^{-4} \leqslant (ra+a^2) \int_k^{\infty} (z-a)^{-4}\, dz = \tfrac{1}{3}(ra-a^2)(k-a)^{-3}.$$

(c) For fixed t let $f(z) = z^r \exp(-z^\beta / a\sqrt{t})$, $z \geqslant 0$. Then $\gamma \equiv (ra\sqrt{t}/\beta)^{1/\beta}$ is the unique maximum of f, so

$$n_j(t) \leqslant \int_j^{j+1} f(z)\, dz, \qquad j \leqslant \lfloor \gamma \rfloor - 1,$$

$$n_j(t) \leqslant \int_{j-1}^{j} f(z)\, dz, \qquad j \geqslant \lceil \gamma \rceil + 1.$$

Also, let $\gamma_+ = \arg\max\{n_{\lfloor\gamma\rfloor}(t), n_{\lceil\gamma\rceil}(t)\}$, and define γ_- similarly. Because f is unimodal we have $n_{\gamma_-}(t) \leqslant f(z)$, $\lfloor\gamma\rfloor \leqslant z \leqslant \lceil\gamma\rceil$, so (provided γ is not an integer) $n_{\gamma_-}(t) \leqslant \int_{\lfloor\gamma\rfloor}^{\lceil\gamma\rceil} f(z)\, dz$. Also, of course, $n_{\gamma_+}(t) \leqslant f(\gamma)$. Thus (whether or not γ is integral),

$$\begin{aligned}
\sum_{j=0}^{\infty} n_j(t) &\leqslant \sum_{j \leqslant \lfloor\gamma\rfloor - 1} \int_j^{j+1} f(z)\, dz + f(\gamma) + \int_{\lfloor\gamma\rfloor}^{\lceil\gamma\rceil} f(z)\, dz \\
&\quad + \sum_{j \geqslant \lceil\gamma\rceil + 1} \int_{j-1}^{j} f(z)\, dz \\
&= \int_0^{\infty} f(z)\, dz + f(\gamma) \\
&= (1/\beta)\Gamma(\lambda)(a\sqrt{t})^{\lambda} + (ra\sqrt{t}/\beta)^{r/\beta} \exp(-r/\beta) \\
&= O(t^{\lambda/2}). \quad \blacksquare
\end{aligned}$$

We are now prepared for our central technical result, Lemma 3. Let ι denote the integers within some finite interval $[l, u]$, $0 \leqslant u \leqslant U$, $l \leqslant u - b$, and Δ_ι the set of policies $\delta \in \Delta_l$ (so $\delta(x) = x + b$, $x \leqslant l$) such that in addition $\delta(x) = x$, $x \geqslant u$. Let $T(\iota)$ be the first period $t \geqslant 1$ with $x_t \in \iota$, conditional on x_0 and $\delta \in \Delta_\iota$. (Note, $T(\iota) > 0$ even if $x_0 \in \iota$.) Finally, for $v \in V_\infty$ define

$$H_\iota v(x) = \max_{\delta \in \Delta_\iota} E\left\{ \sum_{t=0}^{T(\iota)-1} v(y_t) \mid x_0 = x, \delta \right\}, \qquad x < U.$$

LEMMA 3. *If $v(x) = O(|x|^q)$ for a positive integer q, then $H_\iota v(x) = O(|x|^{q+3})$.*

PROOF. *Case* (i): $x < l$. (In this case for every realization $\{D_t\}$ the sum in the expectation is the same for every $\delta \in \Delta_\iota$.) We have

$$\begin{aligned}
|H_\iota v(x)| &< |v(x+b)| \\
&\quad + \sum_{t=1}^{\infty} \sum_{z=-\infty}^{l-1} |v(z+b)| \\
&\quad \times \Pr\{D^{(i)} - ib > x - l,\ 1 < i < t,\ D^{(t)} - tb = x - z\} \\
&< |v(x+b)| + \sum_{t=1}^{\infty} \sum_{z=-\infty}^{l-1} |v(z+b)| \Pr\{D^{(t)} - tb > x - z\} \\
&< |v(x+b)| + \sum_{t=1}^{\infty} \sum_{z=-\infty}^{x} |v(z+b)| \Pr\left\{ \frac{D^{(t)} - t\mu}{\sigma\sqrt{t}} > \sqrt{t}\left(\frac{b-\mu}{\sigma}\right) + \frac{x-z}{\sigma\sqrt{t}} \right\} \\
&\quad + \sum_{t=1}^{\infty} \sum_{z=x+1}^{l-1} |v(z+b)| \Pr\{D^{(t)} - t\mu > t(b-\mu) + (x-l)\}.
\end{aligned} \tag{1}$$

Of the three terms in (1) the first is $|v(x+b)| = O(|x|^q)$. Setting $w = x - z$ and using Lemmas 1 and 2(a), the second term (the first double sum) of (1) is

$$\begin{aligned}
&< \sum_{t=1}^{\infty} \sum_{w=0}^{\infty} |v(x+b-w)| \Pr\left\{ \frac{D^{(t)} - t\mu}{\sigma\sqrt{t}} > \tfrac{1}{2}\left[t^{0.1}\left(\frac{b-\mu}{\sigma}\right) + \frac{w}{\sigma\sqrt{t}} \right] \right\} \\
&< \sum_{t=1}^{\infty} \sum_{w=0}^{\infty} |v(x+b-w)| \left\{ 1 + (C/8t^{1/2}) \left[t^{0.1}\left(\frac{b-\mu}{\sigma}\right) + \frac{w}{\sigma\sqrt{t}} \right]^3 \right\} \\
&\quad \times \left\{ 1 - \Phi\left[\tfrac{1}{2}\left(t^{0.1}\left(\frac{b-\mu}{\sigma}\right) + \frac{w}{\sigma\sqrt{t}} \right) \right] \right\} \\
&< \sum_{t=1}^{\infty} \exp\left[-\tfrac{1}{4} t^{0.1}\left(\frac{b-\mu}{\sigma}\right) \right] \sum_{w=0}^{\infty} |v(x+b-w)| \\
&\quad \times \left\{ 1 + (C/8t^{1/2}) \left[t^{0.1}\left(\frac{b-\mu}{\sigma}\right) + \frac{w}{\sigma\sqrt{t}} \right]^3 \right\} \exp\left(-\tfrac{1}{4} \frac{w}{\sigma\sqrt{t}} \right).
\end{aligned}$$

Now perform binomial expansions of $v(x+b-w) = O(|x-w|^q)$ and of the cubic term, and apply Lemma 2(c), to conclude that the above expression is

$$< \sum_{i=0}^{q} O(|x|^i) \sum_{t=1}^{\infty} t^{[\frac{1}{2}(q-i)+0.3]} \exp\left[-\tfrac{1}{4} t^{0.1}\left(\frac{b-\mu}{\sigma}\right) \right].$$

Applying Lemma 2(c) again, the inner sum is finite for each i, so the entire expression is $O(|x|^q)$. Finally, letting $a = (l - x)/(b - \mu)$ and $k = \lfloor a \rfloor + 2$, the third term of (1) is

$$= \sum_{t=1}^{k} \sum_{z=x+1}^{l-1} |v(z+b)| \Pr\{D^{(t)} - t\mu > t(b-\mu) + (x-l)\}$$

$$+ \sum_{t=k+1}^{\infty} \sum_{z=x+1}^{l-1} |v(z+b)| \Pr\{D^{(t)} - t\mu > t(b-\mu) + (x-l)\}. \tag{2}$$

Each probability is ≤ 1, so the first term of (2) is

$$\leq O\left[k(l-x-1)\max\{|l-1+b|^q, |x+1+b|^q\}\right] = O(|x|^{q+2}).$$

For the second term, the fourth-order Chebyshev inequality (e.g., Thomasian [23, p. 151]) asserts that, for any $\epsilon > 0$,

$$\begin{aligned}\Pr\{D^{(t)} - t\mu > \epsilon\} &\leq (1/\epsilon^4)E(|D^{(t)} - t\mu|^4) = (1/\epsilon^4)E(D^{(t)} - t\mu)^4 \\ &= (1/\epsilon^4)\left[t\mu_{(4)} + 3t(t-1)\sigma^4\right],\end{aligned}$$

where $\mu_{(4)} = E(D-\mu)^4$, the fourth central moment of D. Using $\epsilon = t(b-\mu) + (x-l)$, the second term of (2) is

$$\leq \sum_{t=k+1}^{\infty} \sum_{z=x+1}^{l-1} |v(z+b)| \left[t\mu_{(4)} + 3t(t-1)\sigma^4\right] / \left[t(b-\mu) + (x-l)\right]^4$$

$$\leq 3\left(\frac{\sigma}{b-\mu}\right)^4 \sum_{t=k+1}^{\infty} \sum_{z=x+1}^{l-1} |v(z+b)| \frac{t^2 + rt}{(t-a)^4},$$

where $r = |(\mu_{(4)}/3\sigma^4) - 1|$. Lemma 2(b) implies that this is

$$\leq 3\left(\frac{\sigma}{b-\mu}\right)^4 \sum_{z=x+1}^{l-1} |v(z+b)| \left[\frac{1}{k-a} + \frac{r+2a}{(k-a)^2} + \frac{ra+a^2}{(k-a)^3}\right]$$

$$\leq 3\left(\frac{\sigma}{b-\mu}\right)^4 \sum_{z=x+1}^{l-1} |v(z+b)| \left[1 + (r+2a) + (ra+a^2)\right]$$

$$\leq O\left[(l-1-x)a^2 \max\{|l-1+b|^q, |x+1+b|^q\}\right] = O(|x|^{q+3}).$$

Case (ii): $x \geq u$. In this case

$$H_t v(x) = v(x) + \sum_{j=1}^{\infty} p(x-l+j) H_t v(l-j) + \sum_{j=0}^{x-u-1} p(j) H_t v(x-j) \tag{3}$$

(where the last summation is zero for $x = u$). We show first that the first summation is finite and bounded for $x \geq u$. Indeed, from Case (i) there are constants A_t, B_t such that $|H_t v(l-j)| \leq A_t + B_t |l-j|^{q+3}$, $j \geq 0$, so

$$\sum_{j=1}^{\infty} p(x-l+j) |H_t v(l-j)|$$

$$\leq A_t\left[1 - P(x-l)\right] + B_t \sum_{k=0}^{q+3} \binom{q+3}{k} |l|^{q+3-k} \sum_{j=0}^{\infty} p(j)\left(\left[j-x-l\right]^+\right)^k.$$

This expression is finite (because D has finite moments of all orders) and nonincreasing in x, hence it is bounded. Thus, for some positive constants A and B, (3) yields

$$|H_\iota v(x)| \leqslant A + Bx^q + \sum_{j=0}^{x-u-1} p(j)|H_\iota v(x-j)|. \tag{4}$$

Now define the function $H_\iota^{\#}v$ by means of a renewal-type equation analogous to (4):

$$H_\iota^{\#}v(x) = A + Bx^q + \sum_{j=0}^{x-u-1} p(j)H_\iota^{\#}v(x-j), \qquad x \geqslant u.$$

We show by induction that $H_\iota^{\#}v(x) \geqslant |H_\iota v(x)|$, $x \geqslant u$: To begin, $H_\iota^{\#}v(u) = A + Bu^q \geqslant |H_\iota v(u)|$. Suppose the assertion is true up to some x. Then

$$\begin{aligned} H_\iota^{\#}v(x+1) &= [1-p(0)]^{-1}\left[A + B(x+1)^q + \sum_{j=1}^{x-u} p(j)H_\iota^{\#}v(x+1-j)\right] \\ &\geqslant [1-p(0)]^{-1}\left[A + B(x+1)^q + \sum_{j=1}^{x-u} p(j)|H_\iota v(x+1-j)|\right] \\ &\geqslant |H_\iota v(x+1)|. \end{aligned}$$

Another induction shows that $H_\iota^{\#}v$ is increasing: First,

$$H_\iota^{\#}v(u+1) = [1-p(0)]^{-1}[A + B(u+1)^q] > A + Bu^q = H_\iota^{\#}v(u).$$

Suppose $H_\iota^{\#}v(j) \geqslant H_\iota^{\#}v(j-1)$, $j = u+1, \ldots, x$. Then

$$\begin{aligned} H_\iota^{\#}v(x+1) &= [1-p(0)]^{-1}\left[A + B(x+1)^q + \sum_{j=1}^{x-u} p(j)H_\iota^{\#}v(x+1-j)\right] \\ &> [1-p(0)]^{-1}\left[A + Bx^q + \sum_{j=1}^{x-u-1} p(j)H_\iota^{\#}v(x-j)\right] = H_\iota^{\#}v(x). \end{aligned}$$

Combining these results we have

$$\begin{aligned} |H_\iota v(x)| &\leqslant H_\iota^{\#}v(x) \leqslant [1-p(0)]^{-1}[A + Bx^q + (1-p(0))H_\iota^{\#}v(x-1)] \\ &\leqslant [1-p(0)]^{-1}\sum_{j=u}^{x}(A + Bj^q) = O(x^{q+1}). \end{aligned}$$

Case (iii): $x \in \iota$. Letting $y_0 = \delta(x)$ for the policy realizing $H_\iota v(x)$,

$$H_\iota v(x) = v(y_0) + \sum_{j=1}^{\infty} p(y_0 - l + j)H_\iota v(y_0 - j).$$

As in the proof of Case (ii), the expression is finite for all y_0, $l \leqslant y_0 \leqslant \min\{u+b, U\}$, so $H_\iota v(x)$ is finite. This is all we require, since ι is finite. ∎

The following result is an immediate application of the lemma:

COROLLARY 1. *For any ι satisfying the assumptions of Lemma* 3 $H_\iota v1(x) = O(|x|^3)$ *and* $H_\iota G(x) = O(|x|^{\rho+3})$. *In particular the two functions are finite-valued.*

For any $\delta \in \Delta_L$ define the operator $P[\delta]$ by $P[\delta]v(x) = Ev[\delta(x) - D]$, $x \leqslant U$, $v \in V_\infty$. (Assumpion 2 and $v \in V_\infty$ ensure that the expectation exists. $P[\delta]$ can be thought of as the stochastic transition matrix of $\{x_t | \delta\}$.) Set $L = \bar{y}^\infty - b$ and let $\{(\delta_{0t}, \delta_{1t}, \ldots, \delta_{tt})\}_{t=0}^{\infty}$ be any sequence of $(t+1)$-period policies, $t \geqslant 0$, where $\delta_{st} \in \Delta_L$, $0 \leqslant s \leqslant t$. The proof of Lemma 5 in Federgruen et al. [8], using Lemma 3 and

Corollary 1 above, yields the following result:

COROLLARY 2. *For all* $x \leqslant U$

$$\lim_{t\to\infty} (t+1)^{-1} P[\delta_{0t}] P[\delta_{1t}] \ldots P[\delta_{tt}](H_t v 1 + H_t G)(x) = 0.$$

Next we return to the general case $L < U - b$, and prove an irreducibility result concerning the states in $[L, U]$, assuming $U < \infty$. Let D_- and D_+ denote the first and second smallest possible values of D, i.e., with $p(D) > 0$, respectively; note, $0 \leqslant D_- < D_+ < \infty$, and $D_- < b$. Assume L is chosen sufficiently small that $U - L \geqslant \max\{b, D_- + D_+\}$.

LEMMA 4. *Assume* $U < \infty$. *For every* x' *with* $L \leqslant x' \leqslant U - D_-$ *there exists a policy* $\delta \in \Delta_L$ *such that, for all* x_0 *in* $[L, U - D_-]$, x' *can be reached from* x_0 *under* δ *in a finite number of periods.*

PROOF. Fix x', and define

$$k^+ = \max\{\lceil (x' - L)/D_+ \rceil, \lfloor (U - L)/D_+ \rfloor\},$$

$$X^+ = \{x = L + D_+ k : k = \lceil (x' - L)/D_+ \rceil, \ldots, k^+\},$$

$$x^+ = \max X^+ = L + D_+ k^+, \quad x^- = \min X^+, \quad \text{and} \quad x^{up} = \min\{U - D_-, x^+\}.$$

Observe, X^+ is nonempty, $x^- - D_+ \leqslant x' \leqslant x^-$, $x^{up} \geqslant U - D_+$, and $x^{up} \geqslant x'$.

Define the policy δ as follows:

$$\delta(x) = \begin{cases} x, & x \in X^+ \quad \text{or} \quad x \geqslant x^{up}, \\ x + D_- + 1, & \text{otherwise}, \end{cases}$$

for $L < x \leqslant U$, and $\delta(x) = x + b$, $x \leqslant L$. Observe, $x \leqslant \delta(x) \leqslant x + b$ for all $x \leqslant U$, since $D_- < b$. Also, for $x < U - D_-$ we have $\delta(x) \leqslant x + D_- + 1 \leqslant U$, while for $U - D_- \leqslant x \leqslant U$, we have $x \geqslant x^{up}$, so $\delta(x) = x \leqslant U$. Thus, $\delta \in \Delta_L$.

First, if $x_0 = L < x'$ and $D_0 = D_-$, then $x_1 = L + b - D_- > L$, so it suffices to show x_0 can reach x' for $x_0 > L$. In general, if $L < x_0 < x'$, then $x_0 \notin X^+$ and $x_0 < x^{up}$, so $\delta(x_0) = x_0 + D_- + 1$. Now, suppose $D_t = D_-$ for several periods $t \geqslant 0$. Then $x_{t+1} = x_t + 1$ until $x_t = x'$ when $t = x' - x_0$. And, of course, the result is immediate when $x_0 = x'$. Furthermore, suppose $x_0 > x^{up}$ and $D_0 = D_+$. Clearly, $x^+ \geqslant L + D_+$, and $U - D_- \geqslant L + D_+$ by assumption, so

$$L \leqslant x^{up} - D_+ < x_0 - D_+ = x_1 \leqslant U - D_+ \leqslant x^{up},$$

in particular $x_1 \leqslant x^{up}$. Finally, if $x_0 = x^{up}$ and $D_0 > 0$, then $x_1 < x^{up}$. In sum, it suffices to show that the interval $[L, x']$ can be reached from any $x_0, x' < x_0 < x^{up}$. We distinguish two cases.

(i) $[x_0, x^{up}] \cap X^+$ is nonempty: suppose $x_0 \notin X^+$ and $D_t = D_-$ for several periods $t \geqslant 0$; then $x_{t+1} = x_t + 1$ until some $x_t \in X^+$, so we may presume $x_0 \in X^+$. Now, suppose $D_t = D_+$ for several periods; then $x_{t+1} = x_t - D_+ \in X^+$ until some $x_t = x^-$ If $x' = x^-$ we are done, while if $x' < x^-$, then $x^- \geqslant L + D_+$, and $L \leqslant x_{t+1} = x^- - D_+ \leqslant x'$.

(ii) $[x_0, x^{up}] \cap X^+$ is empty: as in (i), if several $D_t = D_-$, then $x_t = x^{up}$ for $t = x^{up} - x_0$, and if $D_t = D_+$, then $x_{t+1} = x^{up} - D_+ \geqslant L$. If $x' \geqslant x_{t+1}$, we are done, so suppose $x' < x_{t+1}$. Then, $x^- \leqslant x' + D_+ < x^{up}$, so $[x_{t+1}, x^{up}] \cap X^+$ is nonempty, and we may proceed to apply Case (i). ∎

REMARK. If $D_- > 0$, the states $(U - D_-, U]$ are transient under every policy in Δ_L, so the lemma implies that, among states in $[L, U]$ which are positive recurrent under *some* policy, any pair can communicate (perhaps under other policies).

Next we demonstrate an important property of the value-iteration operators defined in §2. Define

$$V = \{v : v \in V_{\rho+3}, v(x) \text{ is nonincreasing, } x < \bar{y}^{\infty}, \text{and } v(\cdot) \text{ is convex}\}.$$

LEMMA 5. *The operators S and Q map V into itself; assuming $L \leqslant \bar{y}^{\infty} - b$, $S_L v = Sv$ and $Q_L v = Qv$, $v \in V$, so S_L and Q_L also map V into itself.*

PROOF. We prove the results for S and S_L; the extension to Q and Q_L is immediate. Choose $v \in V$. $Rv(y) = O(|y|^{\rho+3})$, using $v(y-j) = O(|y-j|^{\rho+3})$ and Assumptions 2 and 3; furthermore, by Assumption 1, Rv is nonincreasing below $\bar{y}^{\infty}$ and convex. Thus, Rv has a global minimum $\bar{y}$, $\bar{y}^{\infty} \leqslant \bar{y} \leqslant U$. It follows that a critcal-number policy achieves the infima defining S and S_L, in particular,

$$S_L v(x) = Sv(x) = \begin{cases} Rv(x+b), & x \leqslant \bar{y} - b, \\ Rv(\bar{y}), & \bar{y} - b < x < \bar{y}, \\ Rv(x), & \bar{y} \leqslant x \leqslant U. \end{cases} \tag{5}$$

Thus, from the properties of Rv and again using Assumption 2, $S_L v = Sv \in V$. ∎

4. Optimality proof (finite capacity). Throughout this section we assume $U < \infty$. The optimality equation for our model can be written

$$gv1 + v = Sv \tag{6}$$

where v is an unknown function, $v \in R_E^U$, and g is an unknown scalar.

THEOREM 1. (a) *There exists a solution (g^*, v^*) to (6), where v^* is convex and has a finite global minimizer $y^* \geqslant \bar{y}^{\infty}$.*

(b) *The critical-number policy δ^* having critical number y^* is strongly optimal, and its average cost is g^*.*

PROOF. (a) Consider a restricted model with $L \leqslant \min\{\bar{y}^{\infty} - b, U - b, U - D_+ - D_-\}$. We shall apply Theorem 1 of Federgruen, Schweitzer and Tijms [9] to this model; to do this requires verifying four assumptions in that paper, which here we label FST1–FST4. (To restate these would require notation well beyond ours, so we merely indicate the reason each is satisfied.) FST1 and FST4 follow immediately from the facts that time is discrete and each $Y_L(x)$ is finite. Part (a) of FST2 is Corollary 1 to Lemma 3 above, and part (b) follows from Lemma 4 and the subsequent remark. Finally, FST3 is immediate from $G(\cdot) \geqslant 0$. It follows from Theorem 1 in [9] that Q_L is a continuous operator, and there exists a compact, convex subset $W' \subset V_{\rho+3}$ with $v0 \in W'$, such that Q_L maps W' into itself.

Now, define $W^* = W' \cap V$. Let $\{v_n\}_{n=1}^{\infty} \to v_{\infty}$ with $v_n \in W^*$, $n \geqslant 1$. Since W' is compact, we have $v_{\infty} \in W'$; also, v_{∞} is convex and nonincreasing below $\bar{y}^{\infty}$, so $v_{\infty} \in W$ implies $v_{\infty} \in V$. Thus, $v_{\infty} \in W^*$, so W^* is closed, and since W' is compact, W^* is compact also (Royden [21, Proposition 2, p. 158]). In addition W^* is convex and $v0 \in W^*$. Finally, the result above and Lemma 5 ensure that Q_L maps W^* into itself.

Using Lemma 5 again, $W^* \subset V$ implies that $Q = Q_L$ on W^*, so Q maps W^* into itself. We now use the proof of Theorem 2 in [9], invoking Tychonoff's fixed-point theorem (cf. Theorem 2.2, p. 414 in Dugundji [6]) to conclude that Q has a fixed point v^* in W^*. Setting $g^* = Sv^*(\bar{y}^{\infty})$, therefore (g^*, v^*) satisfies (6). By construction, v^* is convex; also, v^* has a global minimizer y^* with $y^* \geqslant \bar{y}^{\infty}$, since $v^* \in V$.

(b) We need to show that the restricted model also satisfies Assumption 5′ in [9]. This follows from Lemma 3 with $v = (H_\iota v1 + H_\iota G)$ for $\iota = [L, U]$. Theorem 3 in [9] and $\delta^* \in \Delta_L$ then imply that g^* is the average cost of policy δ^*. (That theorem also ensures that δ^* is optimal for all the restricted models; now we prove the result for the unrestricted case.)

If δ^* is not strongly optimal, there exist a Markov policy $\delta' = (\delta_0', \delta_1', \dots)$ and a state x', such that

$$\liminf_{t\to\infty} t^{-1} E\left\{ \sum_{i=0}^{t-1} G(y_i) \mid x_0 = x', \delta' \right\} < g^*. \tag{7}$$

Also, let $\delta'' = (\delta_n'')_{n=1}^{\infty}$, where δ_n'' is the policy which solves the finite horizon problem with n periods remaining, that is, which achieves the infimum in $S^n v0(x)$ for all $x \leqslant U$. Then, there exist a number $\epsilon > 0$ and a sequence of periods $\{n_k\}_{k=1}^{\infty}$ such that

$$\begin{aligned} n_k g^* &= v^*(x') = S^{n_k} v^*(x') \\ &\leqslant E\left\{ \sum_{i=0}^{n_k-1} G(y_i) \mid x_0 = x', \delta'' \right\} + P[\delta_{n_k}''] \dots P[\delta_1''] v^*(x') \\ &\leqslant E\left\{ \sum_{i=0}^{n_k-1} G(y_i) \mid x_0 = x', \delta' \right\} + P[\delta_{n_k}''] \dots P[\delta_1''] v^*(x') \\ &\leqslant n_k(g^* - \epsilon) + P[\delta_{n_k}''] \dots P[\delta_1''] v^*(x'). \end{aligned} \tag{8}$$

By Lemma 5 and $v0 \in V$ we have $\delta_n'' \in \Delta_L$, $n \geqslant 1$, so the sequence $\{(\delta_{t+1}'', \delta_t'', \dots, \delta_1'')\}_{t=0}^{\infty}$ satisfies the conditions of Corollary 2 above. Also, we can write (6) in the form $v^*(x_t) \leqslant \{G(y_t) - g^* + Ev^*(x_{t+1}) \mid x_t, \delta\}$ for all $t \geqslant 0$, $\delta \in \Delta$, and any x_t, so

$$\begin{aligned} v^*(x) \leqslant{} & E\left\{ \sum_{t=0}^{T(\iota)-1} G(y_t) \mid x_0 = x, \delta \right\} - g^* E\{T(\iota) \mid x_0 = x, \delta\} \\ & + E\{v^*(x_{T(\iota)}) \mid x_0 = x, \delta\}. \end{aligned} \tag{9}$$

((9) is true for any x and ι.) Now, by the definition of $T(\iota)$, $L - b \leqslant x_{T(\iota)} \leqslant U$, so the last term in (9) is bounded in x. Also, for $x \leqslant L$, all policies $\delta \in \Delta_\iota$ agree with δ^*, so $H_\iota v1(x) = E\{T(\iota) \mid x_0 = x, \delta^*\}$ and $H_\iota G(x) = E\{\sum_{t=0}^{T(\iota)-1} G(y_t) \mid x_0 = x, \delta^*\}$. Comparing these expressions with (9), we may write $v^* = O(H_\iota v1 + H_\iota G)$. Thus, Corollary 2 implies $\lim_{k\to\infty} n_k^{-1} P[\delta_{n_k}''] \dots P[\delta_1''] v^*(x') = 0$. Dividing (8) by n_k and letting $k \to \infty$ then yeilds $g^* \leqslant g^* - \epsilon$, a contradiction. ∎

Having shown that some critical-number policy is optimal, we conclude this section with some preliminary results on evaluation of such a policy. Let

$\delta[\bar{y}]$ = critical-number policy with critical number $\bar{y} \leqslant U$.

$g(\bar{y})$ = average cost of $\delta[\bar{y}]$.

Setting $\iota = [\bar{y} - b, \bar{y}]$, note that the expectations defining $H_\iota v1(\bar{y})$ and $H_\iota G(\bar{y})$ are independent of the choice of policy $\delta \in \Delta_\iota$ hence are realized by $\delta[\bar{y}]$, and furthermore $T^* = H_\iota v1(\bar{y})$ is independent of $\bar{y}$. The process $\{y_t\}$ is regenerative under $\delta[\bar{y}]$ with regeneration points at epochs when $y_t = \bar{y}$, T^* is the expected time between such epochs, and $H_\iota G(\bar{y})$ is the expected cost. Both quantities being finite by Corollary 1 to Lemma 3, it follows (cf., e.g., Proposition 5.9 in Ross [20]) that

$$g(\bar{y}) = H_\iota G(\bar{y}) / T^*. \tag{10}$$

This expression can be rewritten in a simpler form: Let

$$q_{kt} = \Pr\{D^{(i)} - ib > 0, i = 1, \dots, t;\ D^{(t)} - tb = k\}, \qquad k, t > 0,$$

$$q_k = \sum_{t=1}^{\infty} q_{kt}, \qquad k > 0, \qquad q_0 = 1.$$

Then

$$T^* = \sum_{k=0}^{\infty} q_k, \qquad H_\iota G(\bar{y}) = \sum_{k=0}^{\infty} q_k G(\bar{y} - k). \tag{11}$$

THEOREM 2. (a) *$g(\bar{y})$ is convex and has a finite minimizer $\bar{y}^*$.*

(b) *y^* minimizes g; setting $\bar{y}^* = y^*$, the policy $\delta^* = \delta[\bar{y}^*]$ is optimal for all problems with $U \geq \bar{y}^*$ (but $U < \infty$); for $U < \bar{y}^*$, $y^* = U$ and $\delta[U]$ is optimal.*

PROOF. (a) Assumption 1 and (11) imply convexity and $\lim_{|\bar{y}| \to \infty} g(\bar{y}) = \infty$, implying that $\bar{y}^*$ is finite.

(b) This follows immediately from part (a) and Theorem 1. ∎

5. Unlimited storage capacity. In this section we assume $U = \infty$. Note that, while Theorem 2 shows that the same $\delta^* = \delta[\bar{y}^*]$ is optimal for all sufficiently large but finite U, this is not enough to prove the result for $U = \infty$. We first show that a solution to (6) exists (Theorem 3 below). Then, we prove that δ^* is optimal for $U = \infty$ under a slight restriction on the policy space (Theorem 4).

LEMMA 7. *Suppose v_1 and v_2 both satisfy the equations*

$$g(\bar{y}^*) + v(x) = G(\delta^*(x)) + P[\delta^*]v(x), \qquad x \leq \bar{y}^*, \quad v(\bar{y}^*) = 0. \tag{12}$$

Then $v_1(x) = v_2(x)$, $x \leq \bar{y}^*$.

PROOF. By the discussion preceding (10), the mean recurrence time of state $x = \bar{y}^* - D_-$ is finite, and every state can reach this one under δ^*. Thus, there is a unique steady-state density $\pi(x)$ under δ^*. Thus, for $x \leq \bar{y}^*$,

$$v_1(x) - v_2(x) = P[\delta^*](v_1 - v_2)(x) = (P[\delta^*])^n (v_1 - v_2)(x)$$

for $n \geq 0$, so

$$v_1(x) - v_2(x) = \lim_{N \to \infty} (N+1)^{-1} \sum_{n=0}^{N} (P[\delta^*])^n (v_1 - v_2)(x)$$

$$= \sum_{i=-\infty}^{\bar{y}^*} \pi(i)(v_1 - v_2)(i).$$

Thus, v_1 and v_2 differ by a constant which must be $v_1(\bar{y}^*) - v_2(\bar{y}^*) = 0$. ∎

THEOREM 3. *There exists a solution (g^*, v^*) to (6), where $g^* = g(\bar{y}^*)$, v^* is nonnegative and convex, and $\bar{y}^*$ is the global minimizer of v^*.*

PROOF. Choose any u, $\bar{y}^* \leq u < \infty$. Using Theorems 1 and 2, we can find v_u such that (g^*, v_u) solves (6) for $U = u$. Note, v_u plus any constant still solves (6) (for $U = u$), so we may specify $v_u(\bar{y}^*) = 0$; thus restricted, v_u also solves (12), and so is uniquely defined for $x \leq \bar{y}^*$, by Lemma 7. Also, by the proof of Theorem 1, $\bar{y}^*$ is the minimizer of v_u, so $g^* + v_u(x) = Rv_u(x)$, $\bar{y}^* \leq x \leq u$. This equation specifies $v_u(x)$ recursively (starting with $x = \bar{y}^* + 1$), hence uniquely. Thus, for $u_2 > u_1 \geq \bar{y}^*$, v_{u_1} and v_{u_2} agree for $x \leq u_1$.

Now define v^* as follows: For each x choose any $u \geq x$ (with $u \geq \bar{y}^*$), and set $v^*(x) = v_u(x)$. By construction δ^* achieves the infimum in $Sv(x)$ for all x, and (g^*, v^*) solves (6). Convexity and nonnegativity of v^* follow immediately from the properties of the v_u. ∎

Next we prove two useful results about $T(\iota)$, v^* and related quantities:

LEMMA 8. *For any interval of integers $\iota = [l, u]$ (finite or infinite) the quantity $E\{v^*(x_{T(\iota)}) \mid x_0 = x, \delta\}$ is bounded over $x \notin \iota$ and $\delta \in \Delta$.*

PROOF. When ι consists of all the integers the result is vacuously true. When ι is finite, $l \leqslant x_{T(\iota)} \leqslant u$ for all x_0 and δ, by the definition of $T(\iota)$, and the result is immediate. When $l > -\infty$ and $u = \infty$, likewise, $l \leqslant x_{T(\iota)} < l + b$ for $x_0 \notin \iota$ (that is, $x_0 < l$) and $\delta \in \Delta$.

Only the case $l = -\infty$, $u < \infty$ remains. Note, if $v^*(x) = v^*(y^*)$, $x \leqslant y^*$, we are done, since the possible values of $v^*(x_{T(\iota)})$ are bounded, as above; otherwise, $\lim_{x \to -\infty} v^*(x) = \infty$, by the convexity of v^*. Thus, if $u > y^*$ we can find $u' \leqslant y^*$ with $v^*(u') \geqslant v^*(u)$; if $u \leqslant y^*$, set $u' = u$. Now, by the definition of $T(\iota)$, $y_{T(\iota)-1} - D_{T(\iota)-1} = x_{T(\iota)} \leqslant u < y_{T(\iota)-1}$, so $u' - D_{T(\iota)-1} \leqslant u - D_{T(\iota)-1} \leqslant x_{T(\iota)} \leqslant u$, and $v^*(x_{T(\iota)}) \leqslant v^*(u' - D_{T(\iota)-1})$. Thus, for all $x \notin \iota$ and $\delta \in \Delta$

$$E\{v^*(x_{T(\iota)}) \mid x_0 = x, \delta\} \leqslant E_D v^*(u' - D),$$

which is finite, since $v^* \in V$. ∎

LEMMA 9. *Let ι be a semi-infinite interval of integers, $\iota = [-\infty, u]$, with $u < \infty$.*
(a) $\lim_{x \to \infty} (1/x) E\{T(\iota) \mid x_0 = x, \delta\} = 1/\mu$, $\delta \in \Delta_\iota$.
(b) $\lim_{x \to \infty} \inf_{\delta \in \Delta} (1/x) E\{T(\iota) \mid x_0 = x, \delta\} = 1/\mu$.
(c) *Now let ι be a semi-infinite interval with $\iota = [l, \infty]$, and $l > -\infty$.*

$$\lim_{x \to -\infty} \inf_{\delta \in \Delta} E\{T(\iota) \mid x_0 = x, \delta\} = \infty.$$

PROOF. (a) Divide $T(\iota)$ into two parts, the time until some $x_t \in \iota$ and the subsequent time, if any, until some $y_t \in \iota$. The expectation of the second part is bounded in $x = x_0$, by the proof of Case (ii) of Lemma 3, and the expectation of the first part is $x/\mu + o(x)$ by the elementary renewal theorem.

(b) If $\delta \notin \Delta_\iota$, then $T(\iota)$ is no less than under some policy in Δ_ι. The result now follows from (a).

(c) The infimum clearly is realized by the policy $\delta(x) = x + b$ for all x. Also, $E\{T(\iota) \mid x_0 = x, \delta\} \geqslant \{T(\iota) \mid x_0 = x, \delta, D_t = 0 \text{ for all } t \geqslant 0\} \geqslant [(l - x)/b] - 1 \to \infty$ as $x \to -\infty$. ∎

To state our optimality result we define Δ_∞ as the class of (stationary and nonstationary) policies, each of which has some finite upper limit beyond which it never orders. (The limit is uniform over time for a given policy, but not over policies.) Recall that Δ is the set of stationary policies.

THEOREM 4. (a) *δ^* is strongly optimal among policies in Δ_∞;*
(b) *δ^* is strongly optimal among policies in Δ.*

PROOF. (a) If not, there is a Markov policy $\delta' \in \Delta_\infty$ and a state x', such that (7) holds. Let u' denote the upper limit for δ', and set $u = \max\{u', \bar{y}^*\}$. We can assume $x' \leqslant u$, since all states $x > u$ are transient under δ'. Then δ' is feasible in the problem with $U = u < \infty$, so (7) contradicts Theorem 1.

(b) Set $\iota = \{\bar{x}\}$ for some fixed state $\bar{x}$, and define

$$r(x) = [v^*(x)]^{-1} \inf_{\delta \in \Delta} E\left\{ \sum_{t=0}^{T(\iota)-1} G(y_t) \mid x_0 = x, \delta \right\},$$

where $r(x) = +\infty$ when $v^*(x) = 0$. We show that $\inf_x r(x) > 0$; the result then follows from Theorem 1 in Robinson [19]. Since $v^* \geqslant 0$ and $G > 0$, we have $r > 0$, and it suffices to show

$$\liminf_{x \to +\infty} r(x) > 0, \tag{13}$$

$$\liminf_{x \to -\infty} r(x) > 0. \tag{14}$$

First, let $\iota' = [-\infty, \bar{x}]$ and note

$$r(x) \geq \left[v^*(x)\right]^{-1} \inf_{\delta \in \Delta} E\left\{ \sum_{t=0}^{T(\iota')-1} G(y_t) \mid x_0 = x, \delta \right\}. \tag{15}$$

Also, (9) and Lemmas 8 and 9(b) imply that, for x sufficiently large,

$$v^*(x) \leq E\left\{ \sum_{t=0}^{T(\iota')-1} G(y_t) \mid x_0 = x, \delta \right\}, \qquad \delta \in \Delta, \tag{16}$$

which immediately yields (13). (15) holds again for $\iota' = [\bar{x}, \infty]$; (9) and Lemmas 8 and 9(c) imply (16) for x sufficiently small, yielding (14). ∎

Note added in proof. As this paper was going to press, we learned about the Chernoff bounds, originally derived by Chernoff, H. (1952). A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations. *Ann. Math. Statist.* **23** 493–509, and discussed in Bienstock, D. (1985). On the Chernoff Bounds. Working Paper OR 139-85, O.R. Center, MIT, Cambridge, MA. This result can be used in place of our Lemma 1, permitting some weakening of Assumptions 2 and 3 and a modest simplification of the proof of Lemma 3.

References

[1] Baker, K., Dixon, M., Magazine, M. and Silver, E. (1978). An Algorithm for the Dynamic Lot-Size Problem with Time-Varying Production Capacity Constraints. *Management Sci.* **24** 1710–1720.

[2] Bitran, G. and Yanasse, H. (1982). Computational Complexity of the Capacitated Lot Size Problem. *Management Sci.* **28** 1174–1186.

[3] Deb, R. and Serfozo, R. (1973). Optimal Control of Batch Service Queues. *Adv. in Appl. Probab.* **5** 362–379.

[4] De Kok, A., Tijms, H. and Van Der Duyn Schouten, F. (1984). Approximations for the Single-Product Production-Inventory Problem with Compound Poisson Demand and Service Level Constraints. *Adv. in Appl. Probab.* **16** 378–402.

[5] Doshi, B., Van Der Duyn Schouten, F. and Talman, J. (1978). A Production-Inventory Control Model with a Mixture of Backorders and Lost-Sales. *Management Sci.* **24** 1078–1086.

[6] Dugundji, J. (1970). *Topology*. Allyn and Bacon, Boston.

[7] Evans, R. (1967). Inventory Control of a Multi-Product System with a Limited Production Resource. *Naval Res. Logist. Quart.* **14** 173–184.

[8] Federgruen, A., Hordijk, A. and Tijms, H. (1979). Denumerable State Semi-Markov Decision Processes with Unbounded Costs, Average Cost Criterion. *Stochastic Process. Appl.* **9** 222–235.

[9] ———, Schweitzer, P. J. and ———. (1983). Denumerable Undiscounted Semi-Markov Decision Processes with Unbounded Rewards. *Math. Oper. Res.* **8** 298–314.

[10] ——— and Zipkin, P. (1986). An Inventory Model with Limited Production Capacity and Uncertain Demands. II. The Discounted-Cost Criterion. *Math. Oper. Res.* **11** 208–215.

[11] Feller, W. (1971). *An Introduction to Probability Theory and its Applications. Vol. II*, 2nd ed. Wiley. New York.

[12] Florian, M. and Klein, M. (1971). Deterministic Production Planning with Concave Costs and Capacity Constraints. *Management Sci.* **18** 12–20.

[13] ———, Lenstra, J. and Rinnooy Kan, A. (1980). Deterministic Production Planning: Algorithms and Complexity. *Management Sci.* **26** 669–679.

[14] Gavish, B. and Graves, S. (1980). A One-Product Production/Inventory Problem under Continuous Review Policy. *Oper. Res.* **28** 1228–1236.

[15] Graves, S. and Keilson, J. (1981). The Compensation Method Applied to a One-Product Production Inventory Problem. *Math. Oper. Res.* **6** 242–262.

[16] Heyman, D. and Sobel, M. (1984). *Stochastic Models in Operations Research*. Vol. II. McGraw-Hill. New York.

[17] Iglehart, D. (1963). Dynamic Programming and Stationary Analysis in Inventory Problems. In H. Scarf, D. Guilford and M. Shelly (Eds.), *Multi-Stage Inventory Models and Techniques*. Stanford University Press, Stanford, Cal.

[18] Karlin, S. (1958). Optimal Inventory Policy for the Arrow-Harris-Marschak Dynamic Model. In K. Arrow et al. (Eds.), *Studies in the Mathematical Theory of Inventory and Production*. Stanford University Press, Stanford, Cal.

[19] Robinson, D. (1980). Optimality Conditions for a Markov Decision Chain with Unbounded Costs. *J. Appl. Probab.* **17** 966–1003.

[20] Ross, S. (1970). *Applied Probability Models with Optimization Applications*. Holden Day, San Francisco.

[21] Royden, H. (1968). *Real Analysis*, 2nd ed. MacMillan, New York.

[22] Scarf, H. (1960). The Optimality of (s, S) Policies in the Dynamic Inventory Problem. In K. Arrow, S. Karlin and P. Suppes (Eds.), *Mathematical Methods in the Social Sciences*. Stanford University Press, Stanford, Cal.

[23] Thomasian, A. (1969). *The Structure of Probability Theory with Applications*. McGraw-Hill, New York.

[24] Tijms, H. (1981). An Algorithm for Denumerable State Semi-Markov Decision Problems with Applications to Controlled Production and Queueing Systems. In R. Hartley et al. (Eds.), *Recent Developments in Markov Decision Theory*. Academic Press, New York.

[25] Veinott, A. (1966). The Status of Mathematical Inventory Theory. *Management Sci.* **12** 745–766.

[26] Wijngaard, J. (1972). An Inventory Problem with Constrained Order Capacity. TH-Report 72-WSK-63, Eindhoven University of Technology.

[27] ———. (1975). Stationary Markovian Decision Problems: Discrete Time, General State Space, Ph.D. Dissertation, Eindhoven University of Technology.

GRADUATE SCHOOL OF BUSINESS, COLUMBIA UNIVERSITY, NEW YORK, NEW YORK 10027