OPTIMAL INVENTORY POLICIES WHEN THE DEMAND DISTRIBUTION IS NOT KNOWN[#]

C. ERIK LARSON[*]

LARS J. OLSON[**]

and

SUNIL SHARMA[***]

UCLA Working Paper No. 631
September 1991

We are grateful to Steven Lippman and Bruce Miller for comments and suggestions. We have also benefitted from discussions with Ken Burdett, Nick Kiefer, Ravi Kumar and Yaw Nyarko.

* School of Business Administration, University of Southern California, Los Angeles, CA 90089-1421.

** Department of Soil and Environmental Sciences, University of California, Riverside, CA 92521-0421, and Visiting Fellow, Department of Agricultural Economics, Cornell University, Ithaca, NY 14853.

*** Department of Economics, University of California, Los Angeles, CA 90024.

OPTIMAL INVENTORY POLICIES WHEN THE
DEMAND DISTRIBUTION IS NOT KNOWN

ABSTRACT

This paper analyzes the stochastic inventory control problem when the demand distribution is not known. In contrast to previous Bayesian inventory models, this paper adopts a non-parametric Bayesian approach that places no restrictions on the prior information about demand, allows for any underlying true (but unknown) demand distribution, and accommodates fixed order costs which make the cost function non-convex. The firm's prior information is characterized by a Dirichlet process prior. As information on the demand distribution accumulates, optimal history-dependent (s,S) rules are shown to converge to an (s,S) rule that is optimal when the underlying demand distribution is known. Two examples are presented. The first illustrates that the optimal policies under learning can differ from those when the distribution is known, even when the prior is such that the initial forecast of the demand distribution coincides with the true distribution. The second example shows that the non-parametric model may perform better than a commonly used (misspecified) parametric model. Comparative dynamic results for the non-parametric model are developed when there is no fixed ordering cost.

OPTIMAL INVENTORY POLICIES WHEN THE DEMAND DISTRIBUTION IS NOT KNOWN

1. INTRODUCTION

Most firms learn about the distribution of demand for their product as they accumulate information over time. Almost all models which incorporate learning adopt a parametric Bayesian approach and assume that the true underlying distribution belongs to some parametric family characterized by a finite number of unknown parameters (Scarf [1959,1960a], Karlin [1960], Iglehart [1964], Azoury and Miller [1984], Azoury [1985], and Lovejoy [1990]). In these models, the firm's information about demand is specified by assuming some conjugate prior on the unknown parameters and updated via Bayes' Rule. The specification of a conjugate family of distributions places restrictions on the prior information that can be accommodated and the true distributions of demand that can be allowed. For example, it is difficult under conjugate family specifications to allow for bi-modal priors or bi-modal true demand distributions. Another key assumption in these parametric Bayesian models is that purchasing costs are linear or convex. It is important to investigate inventory models that encompass both learning and fixed order costs.

This paper adopts a non-parametric Bayesian approach that places no restrictions on the prior information about demand, allows for any underlying true (but unknown) demand distribution, and accommodates fixed order costs which make the cost function non-convex. The firm's information about the demand distribution is characterized by a Dirichlet process prior on the space of distributions (Ferguson [1973]).

Starting with any given Dirichlet process prior, the set of possible posterior distributions is large in the sense that any distribution whose support is included in the support of the measure characterizing the Dirichlet process prior can be approximated as a posterior.

Section 2 develops a dynamic programming formulation of the problem. The state space is defined on the beginning-of-period inventory level and the beginning-of-period forecast of the demand distribution, given information on past demand.[1] In Section 3 we show that for both the finite and infinite planning horizon formulations of the model, a history-dependent (s,S) policy is optimal at each stage. We also prove that as information on the demand distribution accumulates, these history-dependent (s,S) rules converge to the optimal (s,S) rule for the case where the underlying demand distribution is known.

Section 4 provides some illustrative numerical examples. The first example shows that the optimal solution under learning may differ from the solution in the case of a known demand distribution, *even when the prior is such that the initial forecast of demand coincides with the true distribution*. The example also illustrates how the solution under learning varies as confidence in the prior changes. As one would expect, if there is sufficient confidence in a prior that yields a forecast identical to the true distribution, then the learning solution is the same as the solution in the known case. With low confidence in the prior, however, the two solutions may differ. The second example compares the history dependent optimal policies under a misspecified parametric model with those obtained for the non-parametric Bayesian inventory model. For purposes of comparison, the priors are chosen so that the initial forecast of the demand distribution is the same in the two cases. As information about the demand distribution accumulates, the history dependent optimal (s,S) policies in the non-parametric case are closer to the optimal (s,S) policy for known demand distribution than the optimal policies in the former case. As information about demand accumulates, the optimal policies for the non-parametric model eventually approach those for the known case, and give at least as good an approximation as policies in the misspecified parametric model.

Comparative dynamic results for the case where there is no fixed ordering cost are given in Section 5. If the expected demand distributions under two Dirichlet process priors are ordered by first-order stochastic dominance, then the critical numbers are ordered for any common history of demand observations, for all periods-to-go. This extends results obtained by Karlin [1960] and Scarf [1959]. We also show that if in any period the current demand realization is less than (greater than) the minimum (maximum) of those observed to date, the optimal order-to-level S decreases (increases) for the following period. Furthermore, it is possible that if the current demand realization is sufficiently small, then the optimal policy in the following period refrains from ordering additional inventory.

Section 6 offers some concluding remarks. Proofs of all results are given in the Appendix.

2. THE INVENTORY MODEL AND THE DIRICHLET PROCESS

Consider a single commodity periodic review (e.g. at the beginning of each week or month) inventory control problem. The holding and penalty costs are assumed to be linear with h and p the per-unit costs, respectively. These costs are based on the end of period inventory level. There is a fixed order cost K and a per-unit order cost c, where $c < p$. The ordering cost function C is given by

$$C(u) = \begin{cases} K + c\cdot u, & \text{if } u > 0 \\ 0, & \text{if } u = 0. \end{cases}$$

All cost parameters are assumed to be non-negative. The presence of a strictly positive fixed ordering cost makes the ordering cost function non-convex. It is assumed that excess demand is backlogged and there is no lag between ordering and delivery. Future costs are discounted by a single period discount factor β, $0 < \beta < 1$. The true underlying demand distribution, $F^*(\omega)$, is assumed to be stationary.

The firm, not knowing the true demand distribution F^*, takes it to be a random variable defined on the space of distributions on $\mathfrak{R}_+$. Let $\mathcal{B}$ be the σ-algebra of Borel subsets of $\mathfrak{R}_+$, and define $\mathcal{F}$ to be the space of probability measures on $(\mathfrak{R}_+, \mathcal{B})$ with finite variance. $\mathcal{F}$ is endowed with the topology of weak convergence. Let A be the σ-algebra of Borel subsets of $\mathcal{F}$, and let $\mathcal{P}$ be a probability distribution on $(\mathcal{F}, A)$. P denotes a random probability measure chosen according to $\mathcal{P}$, and $\omega_1, .., \omega_n$ denotes a random sample chosen according to P. Define $F(t) = P[0,t]$ so that F(t) is the random probability distribution corresponding to P. Let $E[\cdot]$ denote the expectation operation with respect to $\mathcal{P}$. It is easy to show that E[P] is a probability measure on $(\Re_+,\mathcal{B})$. Let E[F], *the expected distribution function*, be the distribution function corresponding to E[P]. If $\psi(\cdot)$ is any function defined on $\Re_+$, then $E[\psi(z)] = \int_0^\infty \psi(t)dE[F](t)$.

We refer to F as a random probability distribution on $\Re_+$. This means that with probability one, F selects a distribution function with support in $\Re_+$. When we say that F has finite variance, we mean that E[F] has finite variance. The posterior of F after observing $\omega_1,.,\omega_n$ will be denoted by $F|\omega_1,.,\omega_n$ (and sometimes simply by F_n).

It is assumed that P is a Dirichlet process with parameter α, written $P\in D(\alpha)$ or $F\in D(\alpha)$, where α is a finite non-null measure on $(\Re_+,\mathcal{B})$. The Dirichlet process is analyzed in Ferguson [1973]. Ferguson [1974] provides an excellent survey of random distributions.

Definition (Ferguson [1973]). *Let $\alpha(\cdot)$ be a finite non-null measure on $(\Re_+,\mathcal{B})$ and let P be a stochastic process indexed by elements of $\mathcal{B}$. P is a Dirichlet process with parameter α if for every finite measurable partition $(I_1,.,I_m)$ of $\Re_+$, the random vector $(P(I_1),.,P(I_m))$ has a Dirichlet distribution with parameter $(\alpha(I_1),...,\alpha(I_m))$.*

If $P\in D(\alpha)$, then the support of P (in the topology of weak convergence) is the set of all probability measures whose support is contained in the support of α. If $\alpha(\cdot)$ is chosen such that its support is $\Re_+$, then the support of P is $\mathcal{F}$. The one-dimensional marginal distributions of the Dirichlet distribution are Beta distributions so that if $P\in D(\alpha)$ then for every $B\in\mathcal{B}$, P(B) has a Beta distribution with parameters $(\alpha(B), \alpha(\Re_+)-\alpha(B))$. Also, $F(z)$ has a Beta distribution with parameters $(\alpha(z), \alpha(\Re_+)-\alpha(z))$, where $\alpha(z) \equiv \alpha([0,z])$, for all $z \in \Re_+$. Thus $E[F](z) = \alpha(z)/\alpha(\Re_+)$ and the mean of the (expected) prior distribution is $\mu = \int_0^\infty tdE[F](t)$.

When the parameter α of the Dirichlet process has finite support $\{\pi_1,\ldots,\pi_m\}$, one obtains a Dirichlet distribution. In this case the demand takes one of m values, $\omega_1,\ldots,\omega_m$. The underlying demand distribution is discrete with unknown parameter

$$\Pi \in \Delta = \{(\pi_1,\ldots,\pi_m) \mid \pi_i \geq 0, \Sigma\pi_i = 1\},$$

where π_i is the probability that demand is ω_i. The prior distribution on the space Δ of discrete distributions is Dirichlet with parameter $\alpha = (\alpha_1,\ldots,\alpha_m)$. In this case the expected prior distribution is discrete with parameter $(\alpha_1/\Sigma\alpha_i,\ldots,\alpha_m/\Sigma\alpha_i)$. After a demand realization, ω_j, the posterior distribution on Δ is Dirichlet with parameter $(\alpha_1,\ldots,\alpha_j+1,\ldots,\alpha_m)$ and the expected posterior distribution is discrete with parameter $(\alpha_1/\Sigma\alpha_i+1,\ldots,\alpha_j+1/\Sigma\alpha_i+1,\ldots,\alpha_m/\Sigma\alpha_i+1)$ (see DeGroot [1970]).[2]

Under the Dirichlet process the updating of beliefs is special in that information is completely "local". If ω is observed, then the posterior distribution assigns a higher probability to (any subset that contains) ω and uniformly decreases the probability of all subsets that do not contain ω, no matter how far or close they are to ω. Moreover, the posterior random probability distribution is a Dirichlet process.

__Fact 1 (Ferguson [1973])__. *If $F \in D(\alpha)$ and if $\omega_1, \ldots, \omega_n$ is a sample from F, then the posterior distribution of F given $\omega_1, \ldots, \omega_n$, denoted $F|\omega_1, \ldots, \omega_n$, is a Dirichlet process with parameter $\alpha + \Sigma\delta_{\omega i}$, where δ_ω is the measure on $(\Re_+, \mathcal{B})$ that assigns mass one to ω.*

It is easy to see that Fact 1 implies

$$E[F|\omega_1, \ldots, \omega_n](z) = \frac{\alpha(z) + \sum_{i=1}^{n} 1_{[\omega i, \infty)}(z)}{\alpha(\Re_+) + n}$$

$$= \left[\frac{\alpha(\Re_+)}{\alpha(\Re_+) + n}\right] \cdot \frac{\alpha(z)}{\alpha(\Re_+)} + \left[\frac{n}{\alpha(\Re_+) + n}\right] \cdot \frac{\sum_{i=1}^{n} 1_{[\omega i, \infty)}(z)}{n} .$$

Under the Dirichlet process prior assumption the expected posterior distribution is a convex combination of the expected prior distribution and the empirical distribution. $\alpha(\Re_+)$ can be interpreted as a measure of "confidence" in terms of sample size. If $\alpha(\Re_+)$ is large relative to n, then greater weight or "confidence" is placed on the prior. Of course, as the firm accumulates information on demand (and n becomes large) the expected posterior distribution gets closer to the empirical distribution. The Glivenko-Cantelli theorem implies that, in the limit, the firm's expectations about the demand distribution converge to the true demand distribution. It is easy to show that the updating process satisfies the following lemma.

__LEMMA 1.__ *The Bayes mapping $E[F_n] \to E[F_{n+1}|\omega]$ is continuous.*

3. EXISTENCE AND CONVERGENCE OF (s,S) POLICIES UNDER LEARNING

The expected one-period holding and shortage cost given n demand realizations is given by $\int_\Omega L(z,\omega)dE[F_n](\omega)$, where $L(z,\omega) \equiv h\cdot Max[z-\omega,0] + p\cdot Max[\omega-z,0]$. Let $V_T(x,E[F_n])$ be the minimum expected value of discounted costs with T periods to go until the end of the planning horizon when x is the current starting inventory level, $E[F_n]$ is the expected posterior demand distribution given a history of n demand realizations, and an optimal ordering policy is followed in the future. The existence of V_T follows from standard dynamic programming arguments.

THEOREM 2. *There exists an optimal policy that satisfies the following functional equation for $T = 1,\ldots,\infty$:*

$$V_T(x,E[F_n]) = \inf_{u\geq 0} \; C(u)+\int_\Omega (L(x+u,\omega)+\beta V_{T-1}(x+u-\omega,E[F_{n+1}|\omega])dE[F_n](\omega),$$

where $V_0 = 0$. Further, the function V_T is lower-semicontinuous.

Define the post-order inventory level by $z = x + u$. In characterizing the optimal solution it is useful to define the cost functions

$$G_1(z,E[F_n]) = cz+\int_\Omega\{[h\cdot\max[z-\omega,0]+p\cdot\max[\omega-z,0]\}dE[F_n](\omega) \tag{3.1}$$

$$G_T(z,E[F_n]) = cz+\int_\Omega L(z,\omega)dE[F_n](\omega)+\beta\int_\Omega V_{T-1}(z-\omega,E[F_{n+1}|\omega])dE[F_n](\omega). \tag{3.2}$$

Note that G_1 is convex in z and $\lim_{|z|\to\infty} G_1(z,E[F_n]) = \infty$. Let $S_1{}^n$ minimize $G_1(z,E[F_n])$ in z and let $s_1{}^n$ be the smallest value of z such that $G_1(z,E[F_n]) = K + G_1(S_1{}^n,E[F_n])$. Using standard arguments developed by Scarf [1960b], the following facts can be established:

(i) $G_T(z,E[F_n])$ is K-convex in z, (ii) $\lim_{|z|\to\infty} G_T(z,E[F_n]) = \infty$, (iii) $V_T(x,E[F_n])$ is K-convex and continuous in x for $T = 1,\ldots,\infty$, and (iv) there exist scalars (s_T^n, S_T^n) such that S_T^n minimizes $G_T(z,E[F_n])$ in z, where s_T^n is the smallest value of z for which $G_T(z,E[F_n]) = K + G_T(S_T^n,E[F_n])$. Thus, the optimal inventory policy has the following characterization.

THEOREM 3. *There exist (s_T^n, S_T^n) such that the optimal policy satisfies*

$$u_T^* = \begin{cases} S_T^n - x_T & \text{if } x_T < s_T^n \\ 0 & \text{if } x_T \geq s_T^n \end{cases}$$

for $T = 1,\ldots,\infty$.

According to Theorem 3, the optimal policy is an (s,S) inventory rule that varies as expectations change in response to the observed history of demand observations. Given the time-varying nature of optimal policies, it is of substantial interest to determine their long run characteristics. Do optimal policies converge as the number of demand observations increases, and if so, what are the limit policies? An answer to these questions is provided in Theorem 5, below.

Consider the family of T period value functions $\{V_T(x,E[F_n]) | n=1,\ldots\}$, where n is the number of previously observed demand shocks.

LEMMA 4. *The family of value functions $\{V_T(x,E[F_n]) | n=1,\ldots\}$ is equicontinuous at x for all finite T.*

Using Lemma 4 we now prove that as the number of demand observations increases, the history dependent optimal inventory policies under learning converge to policies that are optimal when the true demand distribution is known. Define $u_T^n(x)$ to be the optimal order from x when the current inventory level is x and the estimate of the demand distribution is $E[F_n]$.

THEOREM 5.[3] *For all $T = 1,\ldots,\infty$ (i) $\lim_{n\to\infty} V_T(x,E[F_n]) = V_T^*(x)$ for all x in any finite interval, where V_T^* is the value function for the inventory problem with known demand distribution, F^*, and (ii) for all $\epsilon > 0$ there exists an $N(T,\epsilon)$ such that $0 \le C(u_T^n) + \int[L(x+u_T^n,\omega) + \beta V_{T-1}^*(x+u_T^n-\omega)]dF^* - V_T^*(x) < \epsilon$ for all $n > N(T,\epsilon)$. If $\lim_{n\to\infty} u_T^n(x)$ exists, then $\lim_{n\to\infty} (s_T^n,S_T^n) = (s_T^*,S_T^*)$, where (s_T^*,S_T^*) are optimal for the inventory problem with known demand distribution, F^*.*

This theorem states that once firms observe a sufficient number of demand realizations, their optimal policies under learning will be close to those that are optimal when the true demand distribution is known. It is particularly relevant in situations where information accumulates rapidly, as is the case in many retail and wholesale operations where inventory levels are monitored on a daily or weekly basis.

4. SIMULATIONS

In this section we provide two illustrative numerical simulations. The first example shows that the optimal inventory policy under learning can differ from the optimal policy for a known distribution, *even when* *the expected prior coincides with the true distribution*. This is because it is important to distinguish between distributions on distributions and simple distributions when there is resolution of ncertainty over time. When there is no learning, probabilities of probabilities can be compounded and are equivalent to simple probabilities; however, under learning the two are different (e.g., Kreps [1988]).

This example also demonstrates that optimal (s,S) policies vary with the confidence the firm has in its prior on the demand distribution. If the firm has sufficient confidence in its Dirichlet prior, then the optimal policy under learning is identical to the policy the firm would follow if it knew the demand distribution. With low confidence in the prior, however, the two policies may differ.

The second example compares Bayesian inventory policies for non-parametric Dirichlet and parametric Gamma-Poisson models where the initial expected priors are the same.[4] The example is constructed so that the underlying demand process has a Discrete Uniform distribution, resulting in a misspecified conjugate family for the parametric model. Policies under the non-parametric Dirichlet specification converge more rapidly toward the known demand distribution optimal policy than do the policies under the misspecified Gamma-Poisson model.

EXAMPLE 1

We characterize demand by a Discrete distribution on $\Omega = \{\omega_1, \omega_2, \omega_3\} = \{50, 70, 90\}$ with respective probabilities $p = \{0.7, 0.02, 0.28\}$. The parameters of the model are $h = 0.5$, $p = 2.0$, $c = 1.0$,

$k = 10.0$, $\beta = 0.999$. The problem is solved using the method of successive approximation over a feasible solution set of integer-valued inventory levels. We first solve the problem assuming that the firm knows the demand distribution and then address the unknown case assuming a Dirichlet prior with parameter $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ over the space of discrete distributions on Ω. The parameters are chosen so that the expected prior distribution is the same as the true distribution, but the firm's confidence in the prior is allowed to vary. We examine three cases:

Case 1 : $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (0.7, 0.02, 0.28)$,

Case 2 : $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (1.4, 0.04, 0.56)$,

Case 3 : $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (70.0, 2.0, 28.0)$

In case 1 the firm has very low confidence in its prior [$\sum\alpha_i$ is small], whereas in case 3, the confidence is relatively high [$\sum\alpha_i$ is large].

The T-period optimal policies for $T = 1, 2, 3, 4, 5$ are presented in Table 1. Because the expected prior is the same as the true distribution, the (s,S) policies are the same in all cases for $T = 1$. As it turns out in this example the optimal policies are also identical for $T = 2$. However, for $T \geq 3$, both s and S are larger in case 1 than in the known case, illustrating that with low confidence (and hence greater uncertainty) the firm is liable to overstock. In case 2, only the target level S is higher, while in case 3, the firm has sufficient confidence that its prior guess is correct and the (s,S) policies are identical to the known case. These cases show that even when the expected prior coincides with the actual demand distribution,

TABLE 1

A: Distribution Known

Distribution: Discrete with parameters $\pi=(0.7,0.02,0.28)$

	(s,S) Policies with T Periods to Go					
T=	1	2	3	4	5	6
	(40,50)	(53,100)	(49,90)	(49,90)	(49,90)	(49,90)

B: Distribution Not Known

Case 1.

Prior Distribution: Dirichlet with parameters $\alpha=(0.7,0.02,0.28)$

Expected Prior Distribution: Discrete with parameters $\pi=(0.7,0.02,0.28)$

	(s,S) Policies with T Periods to Go					
T=	1	2	3	4	5	6
	(40,50)	(53,100)	(53,100)	(53,100)	(53,100)	(53,100)

Case 2.

Prior Distribution: Dirichlet with parameters $\alpha=(1.4,0.04,0.56)$

Expected Prior Distribution: Discrete with parameters $\pi=(0.7,0.02,0.28)$

	(s,S) Policies with T Periods to Go					
T=	1	2	3	4	5	6
	(40,50)	(53,100)	(49,100)	(49,100)	(49,100)	(49,100)

Case 3.

Prior: Dirichlet with parameters $\alpha=(70.0,2.0,28.0)$

Expected Prior Distribution: Discrete with parameters $\pi=(0.7,0.02,0.28)$

	(s,S) Policies with T Periods to Go					
T=	1	2	3	4	5	6
	(40,50)	(53,100)	(49,90)	(49,90)	(49,90)	(49,90)

Note: The Discrete distributions are defined on $\Omega=\{50,70,90\}$, and the Dirichlet distributions are defined on the space of Discrete distributions on Ω.

one or both of the critical stocks can differ from those characterizing the optimal inventory policy when the distribution is known.

EXAMPLE 2

Set $h = 2$, $p = 4$, $c = 1$, $k = 4$, $\beta = 0.97$. The true demand distribution is Discrete Uniform on $\Omega = \{\omega_1, \omega_2, \ldots, \omega_7\} = \{0, 1, \ldots, 6\}$. Two cases are considered when the distribution is not known:

Case 1: Poisson(λ) with Gamma(γ, η) conjugate prior,

Case 2: Space of distributions on Ω with Dirichlet(α) conjugate prior.

Case 1 represents the misspecified parametric model and Case 2 the non-parametric model. It would be difficult, if not impossible, to present the solutions for a large number of long demand histories. To simplify the exposition without losing generality, we have chosen to report (s,S) policies for 0, 7, 14, 21, 28, 35, and 105 observations of demand. These observations satisfy the assumption that each demand realization, ω_i, shows up once in the first seven observations, twice in the first fourteen, and so on. This seems reasonable given that the underlying true distribution generating the data is Discrete Uniform on the seven points of support specified by Ω. To make the two cases comparable, the parameters are chosen so that the expected prior distributions are the same (before any observations are taken).

Optimal (s,S) policies for the known and learning models are presented in Table 2. The results show that for both Cases 1 and 2 the (s,S) policies differ substantially from those when the distribution is

TABLE 2

A: Distribution Known

Distribution: Discrete Uniform on Ω

T=	(s,S) Policies with T Periods to Go				
	1	2	3	4	5
	(-4,2)	(0,4)	(0,5)	(0,6)	(0,6)

B: Distribution Not Known

Case 1. Parametric Bayesian Model

Conjugate Prior Family: Poisson(λ) - Gamma(γ,η)

Expected Prior Distribution: Negative Binomial($\gamma,\eta/(1+\eta)$)

(γ,η)	T=	(s,S) Policies with T Periods to Go				
		1	2	3	4	5
25,25		(-5,0)	(-2,1)	(-1,2)	(-1,2)	(-1,3)
46,32		(-5,1)	(-1,2)	(-1,3)	(0,3)	(-1,3)
67,39		(-5,1)	(-1,2)	(0,3)	(-1,4)	(0,4)
88,46		(-4,1)	(-1,3)	(0,4)	(-1,4)	(0,4)
109,53		(-4,1)	(-1,3)	(0,4)	(-1,4)	(0,4)
130,60		(-4,1)	(0,3)	(0,4)	(0,4)	(0,5)
340,130		(-4,2)	(0,4)	(0,5)	(0,5)	(0,5)

Case 2. Non-Parametric Bayesian Model

Conjugate Prior Family: Space of Distributions on Ω - Dirichlet(α)

Expected Prior Distribution: $\alpha(\omega)/\alpha(\Omega)$

$$\alpha = (\alpha_1+j,\alpha_2+j,\alpha_3+j,\alpha_4+j,\alpha_5+j,\alpha_6+j,\alpha_7+j)$$

$$= (.3751+j,\ .3607+j,\ .1803+j,\ .0624+j,\ .0168+j,\ .0037+j,\ .0071+j)$$

j	T=	(s,S) Policies with T Periods to Go				
		1	2	3	4	5
0		(-5,0)	(-2,1)	(-2,2)	(-1,2)	(-1,1)
1		(-5,1)	(0,4)	(0,5)	(0,5)	(0,5)
2		(-4,2)	(0,4)	(0,5)	(0,5)	(0,5)
3		(-4,2)	(0,4)	(0,5)	(0,5)	(0,5)
4		(-4,2)	(0,4)	(0,5)	(0,5)	(0,5)
5		(-4,2)	(0,4)	(0,5)	(0,5)	(0,5)
15		(-4,2)	(0,4)	(0,5)	(0,5)	(0,6)

Note: $\Omega = \{0,1,2,3,4,5,6\}$. The parameters of the Gamma and Dirichlet priors have been chosen so that the expected prior distributions in Cases 1 and 2 are the same.

known, especially for the longer time horizons. After 7 observations however, the (s,S) levels in Case 2 are close to the known case, whereas for Case 1 the policies are quite different even after 14 or 21 observations. As the number of demand observations increase the policies under misspecified parametric learning eventually approach those for the known case, but those under non-parametric learning always give at least as good an approximation.

5. COMPARATIVE DYNAMICS

In this section we assume that there is no fixed ordering cost, i.e., K = 0. We prove two results. The first says that for two Dirichlet process priors, D(a) and D(b) with $a(\Re_+) = b(\Re_+)$, if $E[F](\omega) = a(\omega)/a(\Re_+)$ first order stochastically dominates $E[G](\omega) = b(\omega)/b(\Re+)$ then the optimal inventory level under the first prior is always larger than that under the second prior for any common history of demand realizations. This result complements those obtained by Karlin [1960, Theorems 2 and 2'].

The second result states that if demand in any period is less than the minimum of past demand realizations, then the optimal desired inventory level decreases. Furthermore, it is possible that if demand is low enough, no additional inventory is ordered in the next period. This extends the result of Scarf [1959, Theorem 3] to non-parametric learning.

When it exists, let V'_T denote the derivative of V_T with respect to x. The following lemma leads to the main result.

LEMMA 6. *If E[F] dominates E[G] by first order stochastic dominance then* $V'_T(x,E[F_n]) \le V'_T(x,E[G_n])$. *(If the derivative is not well-defined then the inequality holds for the right hand and left hand derivatives.)*

The main result of this section is stated below.

THEOREM 7. *Assume* $K = 0$.

(i) Consider two Dirichlet process priors, D(a) and D(b), such that $a(R+) = b(R+)$. *For simplicity assume that a and b are continuous with continuous first derivatives.*[5] *Let the expected distributions before any observations are taken be given by* $E[F](z) = a(z)/a(\Re_+)$ *and* $E[G](z) = b(z)/b(\Re_+)$. *If E[F](z) first-order stochastically dominates E[G](z), then the critical numbers satisfy* $S_T^n(F) \ge S_T^n(G)$ *for all n, T and any common history of demand realizations.*

(ii) The critical numbers satisfy

$$S_T^n(E[F_n]) \ge S_T^n(E[F_{n+1} | \omega_{n+1} \le Min(\omega_1, \ldots, \omega_n)])$$

for all n and T. Further, if

$$S_T^n(E[F_n]) > S_T^n(E[F_{n+1} | \omega_{n+1} = 0])$$

then there exists a critical demand level ω^* *such that if* $0 < \omega_{n+1} \le \omega^*$, *the optimal policy is not to order in period n+1.*

6. CONCLUDING REMARKS

This paper uses a two-state dynamic programming approach to analyze the stochastic inventory control problem when the demand distribution is not known. With recent and continuing gains in computing power, it is becomming increasingly possible to calculate optimal inventory rules for models that incorporate non-parametric learning. Further progress could be made if results on the reduction of state space dimensionality, similar to those of Scarf [1960a] and Azoury [1985], are obtained for non-parametric Bayesian inventory models. If this proves difficult or impossible in these models, it would be of interest to develop bounds on the loss from using non-optimal, but simple and readily implementable policies (Lovejoy [1990]).

APPENDIX: PROOFS

Proof of Lemma 1. From Fact 1,

$$E[F_{n+1}|\upsilon](\omega) = (1-a_n)E[F_n](\omega) + a_n 1_{[\upsilon,\infty]}(\omega),$$

where $a_n = \alpha(\Re)/(\alpha(\Re)+n)$. Let $\{E[F_n]_i\} \Rightarrow E[F_n]$, where $\Rightarrow$ denotes convergence in distribution or weak convergence of the associated measures. Then,

$$E[F_{n+1}|\upsilon]_i = (1-a_n)E[F_n]_i + a_n 1_{[\upsilon,\infty]}(\omega) \Rightarrow$$
$$(1-a_n)E[F_n](\omega) + a_n 1_{[\upsilon,\infty]}(\omega) = E[F_{n+1}|\upsilon]. \quad //$$

Proof of Theorem 2. The model satisfies standard continuity assumptions, the cost function is bounded below, and the space of actions over which the cost function is bounded is compact. Under these conditions, the result follows from dynamic programming arguments in Schal [1975]. //

Proof of Theorem 3. The proof follows the classic arguments of Scarf [1960b] and Iglehart [1963].

Proof of Lemma 4. The proof proceeds by induction. Assume $\{V_{T-1}(x,E[F_n])|n=1,\ldots\}$ is an equicontinuous family. Recall that $G_T(z,E[F_n]) = cz + \int_\Omega\{L(z,\omega) + \beta V_{T-1}(z-\omega,E[F_{n+1}|\omega])\}dE[F_n](\omega)$. Since $\{L(z,\omega)|\omega\in\Omega\}$ and $\{V_{T-1}(x,E[F_n]|n=1,\ldots\}$ are equicontinuous families, it is straightforward to show that $\{G_T(z,E[F_n]|n=1,\ldots\}$ is also equicontinuous.

Let (s_T^n, S_T^n) be the optimal policy from $E[F_n]$ with T periods remaining and let z and z′ be the optimal orders from x and x′, respectively. From Theorem 3, it follows that

$$z = \begin{cases} x \text{ if } x > s_T^n \\ S_T^n \text{ if } x \le s_T^n \end{cases}$$

and

$$z' = \begin{cases} x' \text{ if } x' > s_T^n \\ S_T^n \text{ if } x' \le s_T^n. \end{cases}$$

Without loss of generality assume $x' \le x$. We want to show that for any $\epsilon > 0$, there exists a δ_ϵ such that $|x-x'| < \delta_\epsilon$ implies $|V_T(x,\cdot) - V_T(x',\cdot)| < \epsilon$. The following three cases need to be considered.

Case 1: $z = z' = S_T^n$.

In this case, $|V_T(x,\cdot) - V_T(x',\cdot)| = |-cx + G_T(S_T^n, E[F_n]) - (-cx' + G_T(S_T^n, E[F_n]))| = c(x - x')$. Given ϵ choose $\delta_\epsilon = \epsilon/c$. Equicontinuity of the family $\{V_T(x, E[F_n]) | n=1, \ldots\}$ follows directly.

Case 2: $z = x, \quad z' = S_T^n$.

From $x' \le s_T^n \le S_T^n$ it follows that $|z-z'| = |x-S_T^n| \le |x-s_T^n| \le |x-x'|$. Note that

$$V_T(x, E[F_n]) = \begin{cases} K - cx + G_T(S_T^n, E[F_n]) & \text{if } x < s_T^n \\ -cx + G_T(x, E[F_n]) & \text{if } x \ge s_T^n, \end{cases}$$

where $G_T(z, E[F_n])$ is defined above. This gives

$$\begin{aligned} &|V_T(x,\cdot) - V_T(x',\cdot)| \\ &= |-cx + G_T(x,\cdot) - \{K - cx' + G_T(S_T^n,\cdot)\}| \\ &= |-c(x-x') + G_T(x,\cdot) - G_T(s_T^n,\cdot)|. \end{aligned}$$

The last equality follows from the fact that $G_T(s_T^n,\bullet) = K + G_T(S_T^n,\bullet)$ by the K-convexity of G and the definitions of s_T^n and S_T^n. Thus,

$$|V_T(x,\bullet) - V_T(x',\bullet)| = |-c(x-x') + G_T(x,\bullet) - G_T(s_T^n,\bullet)|$$
$$\le c(x-x') + |G_T(x,\bullet) - G_T(s_T^n,\bullet)|.$$

By the equicontinuity of the family $\{G_T(x,\bullet)|n=1,\ldots\}$ there exists a $\delta_{\epsilon/2} > 0$ such that $|x-x'| < \delta_{\epsilon/2}$ implies $|G_T(x,\bullet) - G_T(x',\bullet)| < \epsilon/2$. Then $|x-s_T^n| \le |x-x'| < \delta_{\epsilon/2}$ implies $|G_T(x,\bullet) - G_T(s_T^n,\bullet)| < \epsilon/2$. Define $\delta'_{\epsilon/2} = \epsilon/2c$. From this we get $c(x-x') < \epsilon/2$ whenever $|x-x'| < \delta'_{\epsilon/2}$. Choose $\delta_\epsilon = \text{Min}(\delta_{\epsilon/2},\delta'_{\epsilon/2})$. For this δ_ϵ it can be seen that $|x-x'| < \delta_\epsilon$ implies $|V_T(x,\bullet) - V_T(x',\bullet)| < \epsilon$.

Case 3: $\quad z = x, \quad z' = x'.$

In this case, $|z-z'| = |x-x'|$ and

$$|V_T(x,\bullet) - V_T(x',\bullet)| = |-cx + G_T(x,\bullet) + cx' - G_T(x',\bullet)|$$
$$= |-c(x-x') + G_T(x,\bullet) - G_T(x',\bullet)|$$
$$\le c(x-x') + |G_T(x,\bullet) - G_T(x',\bullet)|.$$

The equicontinuity of the family $\{V_T(x,E[F_n])|n=1,\ldots\}$ follows directly from the fact that $\{G_T(x,E[F_n]|n=1,\ldots)\}$ is an equicontinuous family.

The proof is completed by noting that $\{V_0(x,E[F_n])=0|n=1,\ldots\}$ is an equicontinuous family. //

<u>Proof of Theorem 5.</u> (i) First consider the case of finite T. Let x be in some finite interval $[0,\bar{x}]$. Consider the sequence $\{V_T(x,E[F_n])\}_{n=1}^{\infty}$. $V_T(x,E[F_n])$ is uniformly bounded on $[0,\bar{x}]$ (i.e., there exists an M such that $V_T(x,E[F_n]) \le M$ independent of n). The Ascoli-Arzela theorem (Royden [1988], Theorem 7.40) implies that there exists a subsequence $E[F_{nk}]$ such that $V_T(x,E[F_{nk}])$ converges to a continuous function

$V_T(x,F^*)$ where the convergence is uniform on each compact subset of $[0,\bar{x}]$. Since $E[F_{nk}]$ converges to F^* for all possible subsequences, nk can be taken to be the entire sequence, n. It remains to show that $V_T(x,F^*) = V_T^*(x)$. We need the following preliminary lemma adapted from Hinderer [1970], Lemma 3.3.

<u>Lemma.</u> *If u(x) and v(x) are continuous functions bounded below then*

$$|\inf_x u(x) - \inf_x v(x)| \leq \sup_x |u(x)-v(x)|.$$

<u>Proof of Lemma.</u> Without loss of generality assume that $\inf_x u(x) > \inf_x v(x)$. This implies $\inf u - \inf v = |\inf u - \inf v|$. For any $\epsilon > 0$ such that $\inf u - \inf v \geq \epsilon$ there exists an x for which $u(x) \geq \inf u \geq \inf v + \epsilon \geq v(x)$. This implies $\inf u - \inf v - \epsilon \leq u(x) - v(x) \leq \sup_x |u(x) - v(x)|$, or $|\inf u - \inf v| \leq \sup_x |u(x) - v(x)| + \epsilon$. Since ϵ can be made arbitrarily small the proof is complete. //

The proof of theorem 5 proceeds by induction. Suppose the theorem holds for T-1 and consider $\lim_{n\to\infty} V_T(x,E[F_n])$. By the Lemma,

$$|V_T^*(x) - V_T(x,E[F_n])| \leq \sup_u |C(u)+\int[L(x+u,\omega)+ \beta V_{T-1}^*(x+u-\omega)dF^*(\omega)] - \{C(u)+\int[L(x+u,\omega)+\beta V_{T-1}(x+u-\omega,E[F_{n+1}|\omega])]dE[F_n](\omega)\}|$$

$$= \sup_z |\int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dF^*(\omega) - \{\int[L(z,\omega)+\beta V_{T-1}(z-\omega,E[F_{n+1}|\omega])]dE[F_n](\omega)\}|$$

$$= \sup_z |\int L(z,\omega)dF^*(\omega) - \int L(z,\omega)dE[F_n](\omega) + \beta\int V_{T-1}^*(z-\omega)dF^*(\omega) - \beta\int V_{T-1}^*(z-\omega)dE[F_n](\omega) + \beta\int V_{T-1}^*(z-\omega)dE[F_n](\omega) - \beta\int V_{T-1}(z-\omega,E[F_{n+1}|\omega])dE[F_n](\omega)|$$

$$\le \sup_z |\int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dF^*(\omega) - \int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dE[F_n](\omega)|$$
$$+ \sup_z \beta|\int[V_{T-1}^*(z-\omega) - V_{T-1}(z-\omega,E[F_{n+1}|\omega])]dE[F_n](\omega)|$$
$$\le \sup_z |\int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dF^*(\omega) - \int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dE[F_n](\omega)$$
$$+ \sup_j \sup_z \beta|\int[V_{T-1}^*(z-\omega) - V_{T-1}(z-\omega,E[F_{n+1}|\omega])]dE[F_j](\omega)|, \tag{5.1}$$

where $z = x+u$.

Since $L(z,\omega)+\beta V_{T-1}^*(z-\omega)$ is continuous in ω and $E[F_n]$ converges weakly to F^*,

$$\lim_{n\to\infty} \int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dE[F_n](\omega) = \int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dF^*(\omega),$$

where the convergence is uniform on finite intervals. Thus, the lim and sup operators can be exchanged (Hinderer [1970, remark following Lemma 3.4]) to obtain

$$\lim_{n\to\infty} \sup_z |\int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dF^*(\omega) - \int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dE[F_n](\omega)|$$
$$= \sup_z \lim_{n\to\infty} |\int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dF^*(\omega) - \int[L(z,\omega)+\beta V_{T-1}^*(z-\omega)]dE[F_n](\omega)| = 0. \tag{5.2}$$

Since the period T-1 cost function is uniformly bounded for all n,[6] and $V_{T-1}(z-\omega,E[F_{n+1}|\omega]) \to V_{T-1}^*(z-\omega)$ by the induction hypothesis, the dominated convergence theorem implies

$$\int V_{T-1}(z-\omega,E[F_{n+1}|\omega])dE[F_j](\omega) \to \int V_{T-1}^*(z-\omega)dE[F_j](\omega)$$

which in turn gives

$$\lim_{n\to\infty} \sup_j \sup_z \beta|\int[V_{T-1}^*(z-\omega) - \int V_{T-1}(z-\omega,E[F_{n+1}|\omega])]dE[F_j](\omega)|$$
$$= \sup_j \sup_z \lim_{n\to\infty} \beta|\int[V_{T-1}^*(z-\omega) - \int V_{T-1}(z-\omega,E[F_{n+1}|\omega])]dE[F_j](\omega)| = 0.$$

Combining this with (5.2) shows that (5.1) converges to zero as $n\to\infty$ which proves that $\lim_{n\to\infty} V_T(x,E[F_n]) = V_T^*(x)$. Since the result holds trivially for T=0, the induction argument is complete and the theorem holds for finite T.

Now consider the infinite horizon case. Standard arguments imply:
(1) $V_\infty(x, E[F_n])$ and $V_\infty^*(x)$ are continuous in x, and
(2) $\lim_{T\to\infty} V_T(x, E[F_n]) = V_\infty(x, E[F_n])$ and $\lim_{T\to\infty} V_T^*(x) = V_\infty^*(x)$, where the convergence is uniform in each case. This implies that for all x and $\epsilon > 0$ there exists a T_ϵ such that $|V_T(x, E[F_n]) - V_\infty(x, E[F_n])| < \epsilon$ and $|V_T^*(x) - V_\infty^*(x)| < \epsilon$ for all $T > T_\epsilon$. Combining these two inequalities gives $|V_T(x, E[F_n]) - V_\infty(x, E[F_n]) - V_T^*(x) + V_\infty^*(x)| < 2\epsilon$ for all $T > T_\epsilon$. Taking the limit of this as $n\to\infty$ implies $|V_\infty^*(x) - V_\infty(x, E[F_n])| < 2\epsilon$, since $\lim_{n\to\infty} V_T(x, E[F_n]) = V_T^*(x)$. The proof follows immediately.

(ii) The proof is similar to the proof of Theorem 3.8 in Stokey, Lucas and Prescott [1989]. Define

$W_T^*(x,u) = C(u) + \int_\Omega \{L(x+u,\omega) + \beta V_{T-1}^*(x+u-\omega)\} dF^*(\omega)$, and

$W_T^n(x,u) = C(u) + \int_\Omega \{L(x+u,\omega) + \beta V_{T-1}(x+u-\omega, E[F_{n+1}|\omega])\} dE[F_n](\omega)$.

Note that $V_T^*(x) = \max_u W_T^*(x,u)$ and $V_T(x, E[F_n]) = \max_u W_T^n(x,u)$. Part (i) of the theorem implies that $W_T^n(x,u)$ converges to $W_T^*(x,u)$ where the convergence is uniform on each compact subset $[0,\bar{x}]\times[0,\bar{u}]$. By the principle of optimality

$$0 \le W_T^*(x,u_T^n) - V_T^*(x) \le W_T^*(x,u_T^n) - W_T^*(x,u_T^*) + W_T^n(x,u_T^*) - W_T^n(x,u_T^n)$$
$$\le 2\sup\{|W_T^*(x,u_T^n) - W_T^n(x,u_T^n)|, |W_T^*(x,u_T^*) - W_T^n(x,u_T^*)|\}$$

for all x. Since W_T^n converges to W_T^* uniformly, it follows that for all $\epsilon > 0$ there exists an $N(T,\epsilon)$ such that $0 \le W_T^*(x,u_T^n) - V_T^*(x) < \epsilon$ for all $n > N(T,\epsilon)$. //

Proof of Lemma 6. The proof uses an induction argument. Consider the case when T = 1. The value function is given by

$$V_1(x,E[F_n]) = \inf_{u\geq 0} [c\cdot u + \int_\Omega L(x+u,\omega)dE[F_n](\omega)]$$

$$= \inf_{u\geq 0} [G_1(x+u,E[F_n]) - c\cdot x]$$

where the function G_1 is defined in (3.1). Under the Dirichlet process, G_1 can be rewritten as

$$G_1(z,E[F_n]) = cz + (1-\alpha_n)[h\cdot\int_0^z(z-\omega)dE[F](\omega) + p\cdot\int_z^\infty(\omega-z)dE[F](\omega)]$$

$$+ \alpha_n[h\cdot\frac{1}{n}\sum_{i=1}^{n} \max[z-\omega_i,0] + p\cdot\frac{1}{n}\sum_{i=1}^{n} \max[\omega_i-z,0]]$$

$$= cz + (1-\alpha_n)[h\cdot\int_0^z(z-\omega)dE[F](\omega) + p\cdot\int_z^\infty(\omega-z)dE[F](\omega)]$$

$$+ \alpha_n[h\cdot\frac{1}{n}\sum_{i=1}^{n} (z-\omega_i)1_{(z\geq\omega_i)} + p\cdot\frac{1}{n}\sum_{i=1}^{n} (\omega_i-z)1_{(z<\omega_i)}].$$

As stated earlier, $G_1(z,E[F_n])$ is continuous and convex in z. The derivative of G_1 with respect to z is

$$G'_1(z,E[F_n]) = c + (1-\alpha_n)[h\cdot E[F](z)-p\cdot(1-E[F](z))]$$

$$+ \alpha_n[h\cdot\frac{1}{n}\sum_{i=1}^{n} 1_{(z\geq\omega_i)} + p\cdot\frac{1}{n}\sum_{i=1}^{n} 1_{(z<\omega_i)}]$$

$$= c + h\cdot E[F_n](z) - p\cdot(1-E[F_n](z))$$

$$= c - p + (h+p)\cdot E[F_n](z).$$

This derivative exists except possibly at $\omega_1,\ldots,\omega_n$ and points of discontinuity of the prior; however, at these points the left-hand and right-hand derivatives exist and are bounded. (For simplicity we assume in the sequel that the prior measure has a density.)

It is clear that $G'_1(z,E[F_n])$ is increasing in z, and given p > c we have

$$\lim_{|z|\to\infty} G_1(z,E[F_n]) = \infty.$$

The above discussion implies that a minimum of $G_1(z,E[F_n])$ exists. Let $S_1{}^n(F)$ be the smallest value minimizing $G_1(z,E[F_n])$. The value function and its derivative are given by

$$V_1(x,E[F_n]) = \begin{cases} -c\cdot x + G_1(S_1{}^n,E[F_n]), & x \le S_1{}^n \\ -c\cdot x + G_1(x,E[Fn]), & x \ge S_1{}^n \end{cases}$$

and

$$V'_1(x,E[F_n]) = \begin{cases} -c, & x \le S_1{}^n \\ -c + G'_1(x,E[F_n]), & x \ge S_1{}^n, \end{cases}$$

the derivative being well-defined except possibly at $\omega_1,\ldots,\omega_n$. The continuity and convexity of $G_1(x,E[F_n])$ imply that $V_1(x,E[F_n])$ is continuous and convex in x. Further, $V'_1(x,E[F_n])$ is non-decreasing in x (because if $G'_1(x,E[F_n])$ is well-defined then it is non-negative for $x > S_1{}^n$ by the convexity of G_1).

The T-period value function is defined by

$$V_T(x,E[F_n]) = \inf_{u\ge 0}\ c\cdot u + \int_\Omega \{L(x+u,\omega) + \beta V_{T-1}(x+u-\omega,E[F_{n+1}|\omega])\}dE[F_n](\omega)$$

$$= \inf_{u\ge 0}\ -cx + G_T(x+u,E[F_n]),$$

where $G_T(z,E[F_n])$ is defined in (3.2).

Using standard arguments, it can be established that

(i) $G_T(z,E[F_n]$ is continuous and convex in z

(ii) $\lim_{|z|\to\infty} G_T(z,E[F_n]) = \infty$

(iii) $G'_T(z,E[F_n])$ is increasing in z

and hence that

$$V_T(x,E[F_n]) = \begin{cases} -cx + G_T(S_T^n, E[F_n]), & x \le S_T^n \\ -cx + G_T(x,E[F_n]), & x \ge S_T^n, \end{cases}$$

where S_T^n is the smallest value minimizing $G_T(z,E[F_n])$. Next,

$$V'_T(x,E[F_n]) = \begin{cases} -c, & x \le S_T^n \\ -c + G'_T(x,E[F_n]), & x \ge S_T^n, \end{cases}$$

which is well defined except possibly at $\omega_1,..,\omega_n$ and S_T^n. Clearly $V'_T(x,E[F_n])$ is non-decreasing in x. Finally, $V_T(x,E[F_n])$ is continuous and convex in x.

It is clear from the expression for V'_1 that the result is true for T = 1. We now assume it is true for T-1 and show it is true for T. Consider

$$\begin{aligned} &G'_T(x,E[G_n]) \qquad (6.1) \\ &= c - p + (h+p)E[G_n](x) + \beta\int_\Omega V'_{T-1}(x-\omega,E[G_{n+1}|\omega])dE[G_n](\omega) \\ &= c - p + \int_\Omega\{\beta V'_{T-1}(x-\omega,E[G_{n+1}|\omega]) + (h+p)1_{(\omega\le x)}\}dE[G_n](\omega) \\ &\ge c - p + \int_\Omega\{\beta V'_{T-1}(x-\omega,E[G_{n+1}|\omega]) + (h+p)1_{(\omega\le x)}\}dE[F_n](\omega) \\ &\ge c - p + \int_\Omega\{\beta V'_{T-1}(x-\omega,E[F_{n+1}|\omega]) + (h+p)1_{(\omega\le x)}\}dE[F_n](\omega) \\ &= G'_T(x,E[F_n]), \end{aligned}$$

where the first inequality follows from the fact that $E[F_n]$ stochastically dominates $E[G_n]$ and the integrand is decreasing in ω. The integrand decreases in ω because (i) $V'_{T-1}(z,\cdot)$ increases in z, and (ii) $V'_{T-1}(z,E[G_{n+1}|\omega]) \le V'_{T-1}(z,E[G_{n+1}|\omega'])$ for $\omega < \omega'$, by the induction assumption and the fact that $E[G_{n+1}|\omega']$ stochastically dominates $E[G_{n+1}|\omega]$. The second inequality in (6.1) follows from the induction assumption since $E[F_{n+1}|\omega]$ stochastically dominates $E[G_{n+1}|\omega]$. The critical number $S_T^n(F)$ satisfies $G'_T(x,E[F_n]) = 0$ if the derivative exists, or it is the smallest ω for which the right hand derivative of

$G_T(x, E[F_n])$ is greater than or equal to zero. In either case, the fact that $G'_T(x, E[F_n]) \le G'_T(x, E[G_n])$ enables us to conclude that $S_T^n(F) \ge S_T^n(G)$ (since G'_T is increasing in its first argument).

To complete the proof of the lemma, consider the following three possible cases:

(i) $x \le S_T^n(G)$ which implies $V'_{T-1}(x, E[G_n]) - V'_{T-1}(x, E[F_n])$

$= -c + c = 0,$

(ii) $S_T^n(G) \le x \le S_T^n(F)$ which implies $V'_{T-1}(x, E[G_n]) - V'_{T-1}(x, E[F_n])$

$= -c + G'_T(x, E[G_n]) + c = G'_T(x, E[G_n]) \ge 0,$

(iii) $x \ge S_T^n(F)$ which implies $V'_{T-1}(x, E[G_n]) - V'_{T-1}(x, E[F_n])$

$= G'_T(x, E[G_n]) + G'_T(x, E[F_n]) \ge 0.$

If the derivative is not well-defined at a point x, then the above inequalities can be established for the right hand and left hand derivatives. Hence, the lemma is proved. //

<u>Proof of Theorem 7</u>. (i) The proof follows immediately from Lemma 6.
(ii) Note that $S_T^n(E[F_n])$ satisfies $G'_T(z, E[F_n]) = 0$, or if the derivative is not well defined it is the smallest z such that the right hand derivative of $G_T(z, E[F_n])$ is greater than equal to zero. A similar relation characterizes $S_T^n(E[F_{n+1} | \omega_{n+1} \le \text{Min}\{\omega_1, \ldots, \omega_n\}])$. By Lemma 6, $G'_T(z, E[F_n]) \le G'_T(z, E[F_{n+1} | \omega_{n+1} \le \text{Min}\{\omega_1, \ldots, \omega_n\}])$ because $E[F_n]$ first-order stochastically dominates $E[F_{n+1} | \omega_{n+1} \le \text{Min}\{\omega_1, \ldots, \omega_n\}])$. The result follows because $G'_T(z, \cdot)$ is non-decreasing in z.

<u>Remark</u>. Theorem 7 is not true when there is a fixed ordering cost.

REFERENCES

Azoury, K.S. [1985] "Bayes Solution to Dynamic Inventory Models Under Unknown Demand Distributions," *Management Science*, 31, 1150–1160.

______, and B.L. Miller [1984] "A Comparison of the Optimal Ordering Levels of Bayesian and Non-Bayesian Inventory Models," *Management Science*, 30, 993–1003

Blackwell, D and J.B. MacQueen [1973] "Ferguson Distributions via Pólya Urn Schemes," *Annals of Statistics*, 1, 353–355.

Boylan, E.S. [1969] "Stability Theorems for Solutions to the Optimal Inventory Equation," *J. of Applied Probability*, 6, 211–217.

Dvoretzky, A., J. Kiefer and J. Wolfowitz [1952] "The Inventory Problem: II. Case of Unknown Distributions of Demand," *Econometrica*, 20, 450–466.

Ferguson, T.S. [1973] "A Bayesian Analysis of Some Nonparametric Problems," *Annals of Statistics*, 1, 209–230.

Ferguson, T.S. [1974] "Prior Distributions on Spaces of Probability Measures," *Annals of Statistics*, 2, 615–629.

Hinderer, K. [1970] *Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter*. Berlin: Springer-Verlag.

Iglehart, D.L. [1963] "Optimality of (s,S) Policies in the Infinite Horizon Inventory Problem," *Management Science*, 9, 259–267.

______, [1964] "The Dynamic Inventory Problem with Unknown Demand Distribution," *Management Science*, 10, 429–440.

Karlin, S. [1960] "Dynamic Inventory Policy with Varying Stochastic Demands," *Management Science*, 6, 231–258.

Kreps, D. [1988] *Notes on the Theory of Choice*. Boulder: Westview Press.

Lovejoy, W.S. [1990] "Myopic Policies for Some Inventory Models with Uncertain Demand Distributions," *Management Science*, 36, 724-738.

Porteus, E.L. [1990] "Stochastic Inventory Theory," in *Handbooks in OR and MS*, vol. 2, (ed. by D.P. Heyman and M.J. Sobel). Amsterdam: Elsevier.

Royden, H.L. (1988) *Real Analysis* (3rd edition). New York: Macmillan.

Scarf, H. [1959] "Bayes Solutions to the Statistical Inventory Problem," *Annals of Mathematical Statistics*, 30, 490-508.

______, [1960a] "Some Remarks on Bayes Solutions to the Inventory Problem," *Naval Research Logistic Quarterly*, 7, 591-596.

______, [1960b] "On the Optimality of (s,S) Policies in the Dynamic Inventory Problem," in *Mathematical Methods in the Social Sciences 1959*, (ed. by K.J. Arrow, S. Karlin, and P. Suppes). Stanford: Stanford University Press.

Schal, M. [1975] "Conditions for Optimality in Dynamic Programming and for the Limit of n-Stage Optimal Policies to Be Optimal," *Z. Wahrschein-lichkeitstheorie und Verw. gebiete*, 32, 179-196.

Stokey, N., R. Lucas and E. Prescott [1989] *Recursive Methods in Economic Dynamics*. Cambridge: Harvard University Press.

NOTES

1. A contribution of previous work on parametric Bayesian inventory models is that for some conjugate families the dynamic programming problem can be reduced to a one-dimensional state space, improving computational feasibility. With recent increases in computing power, however, optimal policies can be calculated for inventory problems with higher dimensional state spaces.

2. Blackwell and MacQueen [1973] show that the Dirichlet process can be represented by a generalized Pólya urn scheme.

3. While similar, Boylan's [1969] stability result for solutions to the optimal inventory equation does not apply under Bayesian learning.

4. Scarf's [1960] proof of the optimality of (s,S) inventory policies applies to the parametric Bayesian case as well (e.g., Porteus [1990]).

5. These continuity assumptions simplify the proof but are not essential to the result.

6. The boundedness of the cost function follows from the fact that all distribution functions are assumed to have finite variance.