CENSORED NEWSVENDOR MODEL REVISITED
WITH UNNORMALIZED PROBABILITIES

ALAIN BENSOUSSAN
International Center for Decision and Risk Analysis
School of Management, P.O.Box 830688, SM 30
University of Texas at Dallas, Richardson, TX 75083-0688

METİN ÇAKANYILDİRİM
School of Management, P.O.Box 830688, SM 30
University of Texas at Dallas, Richardson, TX 75083-0688

SURESHE P. SETHI
Center for Intelligent Supply Networks
School of Management, P.O.Box 830688, SM 30
University of Texas at Dallas, Richardson, TX 75083-0688

ABSTRACT. This paper revisits the model of the censored newsvendor presented by Ding, Puterman and Bisi [8]. We analyze that model in an infinite-horizon context by using the interesting concept of unnormalized probabilities. The unnormalized probabilities considerably simplify the dynamic programming equation and facilitate the proof of the existence of an optimal policy. They can also be used to give a simple, alternative proof to Ding et al.’s claim that the myopic order quantity is always less than or equal to the optimal order quantity. Importantly, the concept of unnormalized probabilities can be used to treat other important operations research problems with partial observations.

1. Introduction. The newsvendor problem concerns with the optimization of the inventory level at the beginning of a sales season to meet the demand during the season. When the inventory level is more than the demand during the season, costs are incurred on account of the leftover inventory at the end of the season. Otherwise, some penalty is assessed on the unmet demand. Since the leftover inventory is salvaged at the end of a season in the newsvendor problem, the inventory level at the beginning of the next season is always zero. Thus, the order quantity and the inventory level are always the same. They are used interchangeably in this paper.

Although there is an extensive literature on the newsvendor problem, the unobservability of the unmet demand has only recently been emphasized by Lariviere and Porteus [9], Ding, Puterman and Bisi [8] and Bensoussan, Çakanyildirim and Sethi [3]. In this paper, we consider the multiperiod newsvendor problem of Ding et al., in which the demand in each period is observed fully when it is satisfied from the available inventory. Otherwise, only the event that the demand is larger...
than or equal to the inventory is observed. This problem requires inventory level optimization with censored demand data. It is an example of problems with partial observations (Bensoussan [1] and Monahan [12]). In these problems, the state of the dynamic program (DP) is generally the conditional distribution of the system state. For example both in this paper and in Ding et al., the system state is a demand parameter but the DP state is this parameter’s conditional distribution given the censored demand observations.

We replace the conditional probability of the demand parameter by the so-called unnormalized probability that evolves linearly over time. This replacement provides a simpler but an equivalent DP equation. This simplicity facilitates the proofs of results such as the existence of an optimal policy and the comparison of the optimal and myopic order quantities. It also simplifies the computation of an optimal order quantity. It is fair to say that the efficient computation of the optimal order quantities are critical for an implementation of the model in practice. Moreover, the concept of unnormalized probabilities can be used to analyze many other partially observed problems. Bringing the unnormalized probabilities to the attention of the operations research community is an important contribution of this paper.

Another contribution of our paper relates to the Ding et al.’s claim that the myopic order quantity is always smaller than or equal to the optimal order quantity. Although the claim is true, its proof is found to be erroneous by Lu, Song and Zhu [10], when there are more than two periods in the planning horizon. Lu, Song and Zhu [11] provide another proof for the claim. In this paper, we provide a short, simple proof for Ding et al.’s claim. Moreover, our proof can generalize to other related inventory problems with partial observations.

The plan for this paper is as follows. In the next section we set up the model and obtain a DP equation, whose argument is the conditional probability of the demand parameter. In Section 3, we introduce the important concept of unnormalized probabilities and obtain an equivalent DP equation. We prove the Ding et al.’s claim in Section 4. Some concluding remarks are provided in Section 5.

2. Preliminaries. Let \((\Omega, \mathcal{F}, P)\) be the probability space and let \(n \geq 1\) be the period index. Let \(x_n \geq 0\) denote the demand in period \(n\). The demands are assumed to be independently and identically distributed, and each demand depends on an unknown parameter \(\theta\). Given \(\theta\), the demand has the corresponding density and cumulative probability functions \(f(\cdot|\theta)\) and \(F(\cdot|\theta)\). Let \(\bar{F}(\cdot|\theta) = 1 - F(\cdot|\theta)\).

In this paper, we prove the existence of a feedback ordering policy under the assumption that the mean demand is finite. Strictly speaking, we assume that there exists a sufficiently large constant \(M\) such that

\[
\int xf(x|\theta)dx \leq M\theta < \infty
\]  

(1)

for each \(\theta\). The assumption of a finite demand mean is customary in the inventory literature, and it is used to keep the inventory related costs finite. In applications where the unknown parameter \(\theta\) is the mean of the demand distribution, this assumption holds trivially with \(M = 1\). For brevity, we omit the limits of an integral if the integral is taken over \([0, \infty)\) as in (1).

The inventory available to satisfy the demand \(x_n\), or a part thereof, is called \(y_n\). We can think of \(y_n\) to be the inventory after the \(n\)th period order arrives but before the \(n\)th period demand \(x_n\) materializes. We let the sales \(z_n\) be given by

\[
z_n := \min\{x_n, y_n\}.
\]  

(2)
When \(x_n > y_n\), the inventory is not sufficient to meet the demand in period \(n\). In that case, the amount of sales is \(y_n\) and \(x_n - y_n\) is the unmet demand. When the demand is not met, the magnitude of the unmet demand is not observed by the inventory manager (IM). Indeed, the IM observes only the sales. Let \(Z_n\) be the sigma algebra generated by the sales, i.e., \(Z_n := \sigma(z_1, \ldots, z_n)\). Thus, \(Z_n\) is the information available to the IM at the end of period \(n\). Since the IM decides on \(y_n\) at the beginning of period \(n\), \(y_n\) is \(Z_{n-1}\) measurable. In our partially observed demand model when \(x_n \geq y_n\), \(x_n\) is not \(Z_n\) measurable. The IM decides on \(y_n\) at the beginning of period \(n\) after observing the previous sales \(\{z_j : j \leq n - 1\}\).

Let the function \(L(x, y)\), which depends on the demand \(x\) and the available inventory level \(y\) ordered to meet the demand, denote the one-period cost function:

\[
L(x, y) = \begin{cases} 
  cy - h(y - x) & \text{if } x \leq y \\
  cy + b(x - y) & \text{if } y \leq x 
\end{cases}
\]

with \(0 \leq h < c < b\). Costs \(h\) and \(c\) can be interpreted as the salvage value per unit and the ordering cost per unit. Note that leftover inventory is salvaged at the end of every period. Since the demand \(x \geq 0\), \(L(x, y)\) decreases as \(y\) increases for \(y \leq 0\). Thus, we can restrict \(y\) to \(0, \infty\) in our minimization problem. The cost \(b\) is the backorder (resp. lost sales) cost per unit in a backorder (resp. lost sales) case.

With the discount factor \(0 < \alpha < 1\) and with \(y\) defining the sequence of inventory levels \(y = \{y_1, y_2, \ldots \}\), our objective is to minimize

\[
J(y) := \sum_{n=1}^{\infty} \alpha^{n-1} E L(x_n, y_n). \tag{4}
\]

Since the leftover inventory is always salvaged and we require \(y_n \geq 0\), the inventory level at the beginning of each period can be taken to be zero whether or not \(y_n - x_n > 0\). This, however, implies that \(y_{n+1}\) is independent of \(y_n\). In other words, periods are separable in terms of inventory, so there is no inventory state equation in our formulation. On the other hand, periods are coupled by the update of the demand distribution, whose evolution over time is studied next.

Let \(\pi_n\) be the distribution of the parameter \(\theta\) in period \(n\) before observing \(z_n\). That is, the probability density function of \(x_n\) is \(\int f(\cdot | \theta) \pi_n(\theta) \, d\theta\) when \(\pi_n\) is given. We start with a known \(\pi_1\) and obtain the evolution of the process \(\{\pi_n : n \geq 1\}\):

\[
\pi_{n+1}(\theta) = \mathbb{1}_{x_n < y_n} \frac{f(x_n | \theta) \pi_n(\theta)}{\int f(x_n | \theta') \pi_n(\theta') \, d\theta'} + \mathbb{1}_{x_n \geq y_n} \frac{\int_{y_n}^{\infty} f(\xi | \theta) \pi_n(\theta) \, d\xi}{\int_{y_n}^{\infty} f(\xi | \theta') \pi_n(\theta') \, d\xi \, d\theta'}, \tag{5}
\]

This state evolution equation is equivalent to equation (5) in Ding et al. [8].

From (4), with initial \(\pi_1 = \pi\) we define the value function \(V(\pi)\) as follows:

\[
V(\pi) := \inf_y J(y).
\]

By the optimality principle,

\[
V(\pi_n) = \inf_y \left[ \int \int L(x, y) f(x | \theta) \pi_n(\theta) \, d\theta \, dx + \alpha EV(\pi_{n+1}) \right]. \tag{6}
\]
where from (5) we have
\[
EV(\pi) = \int_0^y V \left( \frac{f(x|\cdot)\pi(\cdot)}{\int f(x|\theta')\pi(\theta')d\theta'} \right) \int f(x|\theta)\pi(\theta)d\theta dx \\
+ V \left( \frac{\int f(x|\cdot)\pi(\cdot)d\xi}{\int \int f(x|\theta')\pi(\theta')d\theta' d\xi} \right) \int_y^\infty \int f(x|\theta)\pi(\theta)d\theta dx.
\] (7)

Combining (6) and (7), we obtain the DP equation
\[
V(\pi) = \inf_y \left\{ \int \int L(x, y)f(x|\theta)\pi(\theta)d\theta dx \\
+ \alpha \int_0^y V \left( \frac{\int f(x|\cdot)\pi(\cdot)}{\int f(x|\theta')\pi(\theta')d\theta'} \right) \int f(x|\theta)\pi(\theta)d\theta dx \\
+ \alpha V \left( \frac{\int f(x|\cdot)\pi(\cdot)d\xi}{\int \int f(x|\theta')\pi(\theta')d\theta' d\xi} \right) \int_y^\infty \int f(x|\theta)\pi(\theta)d\theta dx \right\}.
\] (8)

Our DP equation corresponds to equation (8) in Ding et al. [8], but we keep ours more explicit by writing it in terms of only two functions \(f\) and \(L\).

3. Linearization of the State Transition with Unnormalized Probabilities. In this section, we simplify the highly nonlinear evolution equation (5) of \(\pi_n\) by constructing a Zakai-type evolution equation (Zakai [13]) for the unnormalized probabilities. We define the unnormalized probabilities recursively. Starting with \(\rho_1 = \pi_1\), let
\[
\rho_{n+1}(\theta) := 1_{x_n<y_n}f(x_n|\theta)\rho_n(\theta) + 1_{x_n\geq y_n} \int_{y_n}^\infty f(x|\theta)d\theta \rho_n(\theta).
\] (9)

This state equation is linear in \(\rho\). We also define a normalization factor \(\lambda_n\) in each period:
\[
\lambda_n := \int \rho_n(\theta)d\theta.
\] (10)

Since \(\pi_n\) integrates to 1 while \(\rho_n\) integrates to \(\lambda_n\), \(\pi_n\) (resp. \(\rho_n\)) is called the normalized (resp. unnormalized) probability. The normalization factor \(\lambda_n\) allows us to obtain \(\pi_n\) from \(\rho_n\) and vice versa. Indeed, \(\pi_n(\theta) = \rho_n(\theta)/\lambda_n\). Furthermore, we define a new value function
\[
W(\rho) := V \left( \frac{\rho}{\int \rho(\theta')d\theta'} \right) \int \rho(\theta')d\theta'.
\] (11)

Note that \(W\) has unnormalized probability \(\rho\) as an argument.

We have defined \(\rho_n\) in (9) without any reference to \(\pi_n\) in (5), although the equations (9) and (5) look similar. Based on this similarity, the next theorem relates \(\rho_n\) and \(\pi_n\). It also provides a DP equation for \(W\), which does not involve \(V\). From this equation, \(W\) can be found out directly. The proof of the theorem is presented in the appendix.

Theorem 3.1. i) Each unnormalized probability \(\rho_n\) in (9) is related to the normalized probability \(\pi_n\) in (5) uniquely by
\[
\pi_n(\theta) = \frac{\rho_n(\theta)}{\int \rho_n(\theta')d\theta'}.
\] (12)
ii) $W$ given by (11) solves the DP equation
\[
W(\rho) = \inf_y \left\{ \int \int L(x, y) f(x|\theta) \rho(\theta) d\theta dx + \alpha \int_0^y W(\rho(\cdot)f(x|\cdot)) dx + \alpha W(\rho(\cdot)\bar{F}(y|\cdot)) \right\}. 
\] (13)

iii) The solution $y$ of (8) and the solution $y$ of (13) are the same provided that $\pi$ in (8) and $\rho$ in (13) satisfy $\pi = \rho / \int \rho(\theta) d\theta$.

It is worth comparing the DP equations (8) and (13). A glance at these equations reveals that (13) is simpler and shorter than (8). Moreover, the update of $\rho$ in (9) is linear while the update of $\pi$ in (5) is nonlinear. As a result, (13) does not have a denominator which involves the control variable $y$. Thus, the derivative of (13) with respect to $y$ is easier to obtain than the derivative of (8). Hence, finding the minimizer $y$ is easier with (13) than it is with (8).

Theorem 3.1.iii) states that (13) and (8) are equivalent as far as their solutions are concerned. Thus, the optimal inventory levels which minimize (4) can be found by solving (13). This is a very important result as it enables us to work with the easy DP equation (13) by ignoring the difficult DP equation (8) in the remainder of this paper. A major contribution of this paper is in developing a simpler but equivalent DP framework with the unnormalized probabilities. This significantly simplifies the forthcoming proofs as well as the computation of optimal order quantities.

The simplicity of (13) is due to working with unnormalized probabilities as opposed to normalized probabilities. We expect that the concept of unnormalized probabilities would simplify the analysis of other partially observed inventory models. Consequently, unnormalized probabilities could have far reaching applications than we present in this paper.

4. Existence and Comparison Results. The simplicity of (13) facilitates the proof of the existence of a solution $W$ and of a feedback solution $y^*(\rho)$ that minimizes the right-hand side of (13). These are the subject of Theorem 4.1 below. The proof of this theorem is in the appendix. Briefly speaking, the proof works with the maps $T_y(W)$ and $T(W)$ defined as
\[
T_y(W)(\rho) := \int \int L(x, y) f(x|\theta) \rho(\theta) d\theta dx + \alpha \int_0^y W(\rho(\cdot)f(x|\cdot)) dx + \alpha W(\rho(\cdot)\bar{F}(y|\cdot)),
\]
\[
T(W) := \inf_y T_y(W).
\]

We first show that map $T(W)$ is a contraction, and as a consequence, we obtain Theorem 4.1.i. Then, in addition, we establish lower semicontinuity of $W$, the unique solution of (13), and we construct an upper bound for optimal $y$. These yield Theorem 4.1.ii.

**Theorem 4.1.** i) There exists a unique value function $W(\cdot)$ which solves the DP equation (13).

ii) There exists an optimal feedback solution $y^*(\rho)$.

An important issue is to compare the myopic order quantity to the optimal order quantity that can be found from the DP equation (13). The claim of Ding et al. [8]
is that the myopic order quantity is smaller than or equal to the optimal one. To formally compare the optimal and myopic solutions, let

\[ y^*(\rho) = \arg \min_y \left\{ \int \int L(x, y) f(x|\theta) \rho(\theta) d\theta dx \right. \]

\[ + \alpha \int_0^y W(\rho(\cdot)f(x|\cdot)) dx + \alpha W(\rho(\cdot)F(y|\cdot)) \right\}, \quad (14) \]

\[ y^M(\rho) = \arg \min_y \left\{ \int \int L(x, y) f(x|\theta) \rho(\theta) d\theta dx \right\}, \quad (15) \]

where we have replaced “inf” with “min” in view of Theorem 4.1 and where \( y^*(\rho) \) and \( y^M(\rho) \) denote the optimal and myopic inventory levels, respectively.

The main claim in Ding et al. [8] is \( y^M(\rho) \leq y^*(\rho) \). Their proof was found to be erroneous by Lu et al. [10], when there are more than two periods in the horizon. Note that \( y^*(\rho) \) in (14) solves the infinite horizon problem. A detailed account of the errors in Ding et al. can be found in Lu et al. [10]. In a nutshell, the proof in Ding et al. ended up ignoring some terms while taking the derivatives of the value function. This ignorance could perhaps be caused by the very lengthy expressions in Ding et al. who did not use unnormalized probabilities.

For the same claim, Lu, Song and Zhu [11] provide a proof. Their sample-path proof is based on a key observation that the derivative of the value function can be interpreted as the value function of some policy, which is not optimal. While this observation completes the proof, the identification of the policy which appears in the derivative is itself quite complex. Like Ding et al., Lu, Song and Zhu [11] do not use unnormalized probabilities either. For all these reasons, the proof in Lu et al. [11] is also intricate. Moreover, one does not see the underlying reason for the key observation. Therefore, the generalizations of their proof to other models is not obvious.

Instead, one can use unnormalized probabilities to provide a much simpler proof of the claim. The next theorem is devoted to this claim.

**Theorem 4.2.** The optimal inventory level is greater than or equal to the myopic inventory level, i.e., \( y^*(\rho) \geq y^M(\rho) \) for every \( \rho \).

In this paper, we choose to work with an infinite horizon problem to keep the notation brief and to emphasize the concept of unnormalized probabilities. However, the proof of Theorem 4.2 uses iteration \( W^{n+1} = T(W^n) \). The value function \( W^n \) can be interpreted as the optimal cost to-go when there are \( n \) periods until the end of the horizon. Then our proof immediately implies that the claim of Theorem 4.2 holds also in finite horizon settings.

5. **Concluding Remarks.** In this paper, we revisit the censored demand model of Ding et al. [8] with unnormalized probabilities. These probabilities result in a simple but equivalent DP equation – to the one presented by Ding et al. The simplicity of our DP equation has facilitated the arguments for the existence of an optimal feedback policy. In addition on account of unnormalized probabilities, it results in a short and simple proof of the claim made by Ding et al. that the optimal order quantity is larger than or equal to the myopic order quantity. This provides a closure to the stream of literature beginning with Ding et al. and dealing with comparison of optimal and myopic policies. Moreover, our simple proof is easily
transferable to other situations. For example, with censored Markovian demands in Bensoussan et al. [3], we obtained an analogous result.

Unnormalized probabilities are also shown to be useful in proving the sufficiency of some statistics when demands are exponentially distributed. Namely, the number of stock-out periods and the cumulative sales by period $n$ are shown to be sufficient to describe the distribution of the parameter $\theta$ at the end of period $n$. This sufficiency in turn facilitates the proof of consistency of the Bayesian posteriors $\pi_n$, i.e., the posteriors converge to the true parameter value $\theta_0$ in the limit, regardless of the chosen prior. Convergence both of Bayesian posteriors and of posteriors obtained by an adaptive control scheme are presented in Bensoussan et al. [5]. This consistency result frees the inventory manager from worrying about selecting an appropriate prior, which in most cases is a guess, and lets him focus on inventory management.

The unnormalized probabilities were introduced by Bensoussan, ÇakanyILDIRM and Sethi [2, 3, 4] in the context of partially observed inventories. While they were first developed by Zakai [13] in the context of optimal filtering problems involving Wiener processes, they do not directly apply to inventory problems under consideration. The unnormalized probability approach considerably simplifies the DP formulation and in general leads to relatively simpler proofs. The approach is useful for treating a large variety of operations research problems with partial observations.
Appendix. Proof of Theorem 3.1: i) Since $\rho_1 = \pi_1$ and $\rho_1$ integrates to 1, so we trivially have (12). Now we can prove (12) with an induction on $n$. Suppose that (12) holds for $n$. From (9) and (10), we have

$$\lambda_{n+1} = \int \rho_{n+1}(\theta) d\theta = \mathbb{1}_{x_n<y_n} \int f(x_n|\theta) \rho_n(\theta) d\theta + \mathbb{1}_{x_n>y_n} \int \int_1^\infty f(x|\theta) dx \rho_n(\theta) d\theta.$$

Then,

$$\frac{\rho_{n+1}(\theta)}{\lambda_{n+1}} = \mathbb{1}_{x_n<y_n} \frac{f(x_n|\theta) \rho_n(\theta)}{\int f(x_n|\theta) \rho_n(\theta) d\theta} + \mathbb{1}_{x_n>y_n} \frac{\int_1^\infty f(x|\theta) dx \rho_n(\theta) d\theta}{\int_1^\infty f(x|\theta) dx \rho_n(\theta) d\theta}$$

$$= \mathbb{1}_{x_n<y_n} \frac{f(x_n|\theta) \lambda_n \pi_n(\theta)}{\int f(x_n|\theta) \lambda_n \pi_n(\theta) d\theta} + \mathbb{1}_{x_n>y_n} \frac{\int_1^\infty f(x|\theta) d\lambda_n \pi_n(\theta) d\theta}{\int_1^\infty f(x|\theta) d\lambda_n \pi_n(\theta) d\theta}$$

by inserting $\lambda_n \pi_n$ for $\rho_n$;

$$= \pi_{n+1}(\theta)$$

This equality establishes (12) with index $n+1$ and completes the proof of i).

ii) We start by inserting (8) into (11):

$$W(\rho) = \left( \int \rho(\theta') d\theta' \right) \inf_y \left\{ \int \int L(x,y) f(x|\theta) \left( \frac{\rho(\theta)}{\int \rho(\theta') d\theta'} \right) d\theta dx \right. \right.$$

$$+ \alpha \int_0^y V \left( \frac{f(x|\cdot)}{\int f(x|\theta') (\rho(\cdot)/\int \rho(\theta') d\theta')} \right) \left( \int f(x|\theta) \left( \frac{\rho(\theta)}{\int \rho(\theta') d\theta'} \right) d\theta dx \right.$$

$$+ \alpha V \left( \frac{\int_1^\infty f(x|\theta) (\rho(\cdot)/\int \rho(\theta') d\theta') d\theta}{\int_1^\infty f(x|\theta') (\rho(\cdot)/\int \rho(\theta') d\theta') d\theta} \right) \right\} \left( \int \int \right.$$

$$= \inf_y \left\{ \int \int L(x,y) f(x|\theta) \rho(\theta) d\theta dx \right. \right.$$  

$$+ \alpha \int_0^y V \left( \frac{f(x|\cdot)}{\int f(x|\theta') \rho(\theta') d\theta'} \right) \int f(x|\theta) \rho(\theta) d\theta dx \right.$$  

$$+ \alpha V \left( \frac{\int_1^\infty f(x|\theta) \rho(\cdot) d\theta}{\int_1^\infty f(x|\theta') \rho(\cdot) d\theta} \right) \int_1^\infty \int f(x|\theta) \rho(\theta) d\theta dx \right\} \right.$$  

$$= \inf_y \left\{ \int \int L(x,y) f(x|\theta) \rho(\theta) d\theta dx \right.$$  

$$+ \alpha \int_0^y W \left( f(x|\cdot) \rho(\cdot) dx \right) + \alpha W \left( \int_1^\infty f(x|\cdot) \rho(\cdot) dx \right) \right\} \right.$$  

where the last equality is obtained by using (11) and replacing $V$ with $W$. The last equality is exactly (13), so ii) is proved.
iii) We note that (16) is the same as the right-hand side of (8), while (17) is the same as the right-hand side of (13). Then iii) follows immediately. □

Proof of Theorem 4.1: We need to define functional spaces and appropriate norms in these spaces. Let

$$\mathcal{H} := \left\{ \rho \in L^1(\mathbb{R}^+) : \int_0^\infty x|\rho(x)|dx < \infty \right\},$$

where $L^1(\mathbb{R}^+)$ is the space of integrable functions. We are ultimately interested in unnormalized probabilities, which are always nonnegative. Accommodate these in $\mathcal{H}^+ := \{ \rho \in \mathcal{H} | \rho \geq 0 \}$, where we note that $\mathcal{H}^+$ is a closed subset of $\mathcal{H}$ with the norm

$$||\rho|| := \int_0^\infty |\rho(x)|dx + \int_0^\infty x|\rho(x)|dx. \quad (18)$$

We define the space $\mathcal{B}$ of functions $\phi$:

$$\mathcal{B} = \left\{ \phi(\rho) : \mathcal{H}^+ \rightarrow \mathbb{R} \bigg| \sup_{\rho \geq 0} |\phi(\rho)| ||\rho|| < \infty \right\}$$

with the norm

$$||\phi||_\mathcal{B} = \sup_{\rho \in \mathcal{H}^+} \frac{|\phi(\rho)|}{||\rho||}. \quad (19)$$

We prove the existence of a feedback ordering policy under the assumption that the mean demand is finite; see (1).

i) We first establish that $T$ is a contraction mapping: $||T(W) - T(\tilde{W})||_\mathcal{B} \leq \alpha||W - \tilde{W}||_\mathcal{B}$ for $W, \tilde{W} \in \mathcal{B}$. Note that

$$T_y(W)(\rho) - T_y(\tilde{W})(\rho) = \alpha \int_0^y W(\rho(\cdot)f(x|\cdot)) - \tilde{W}(\rho(\cdot)f(x|\cdot))dx$$

$$+ \alpha \left( W(\rho(\cdot)\tilde{F}(y|\cdot)) - \tilde{W}(\rho(\cdot)\tilde{F}(y|\cdot)) \right),$$

where $T_y(W)$ equals the terms inside the curly brackets on the right-hand side of (13). Then it follows that

$$|T(W)(\rho) - T(\tilde{W})(\rho)| \leq |T_y(W)(\rho) - T_y(\tilde{W})(\rho)|$$

$$\leq \alpha||W - \tilde{W}||_\mathcal{B} \int_0^y ||\rho(\cdot)f(x|\cdot)||dx$$

$$+ \alpha||W - \tilde{W}||_\mathcal{B}||\rho(\cdot)\tilde{F}(y)|| \quad (20)$$

We now examine the terms inside the curly brackets on the right-hand side of (20):

$$\int_0^y ||\rho(\cdot)f(x|\cdot)||dx + ||\rho(\cdot)\tilde{F}(y)||$$

$$= \int_0^y \left\{ \int \rho(\xi)f(x|\xi)d\xi + \int \xi\rho(\xi)f(x|\xi)d\xi \right\}dx$$

$$+ \int \rho(\xi)\tilde{F}(y|\xi)d\xi + \int \xi\rho(\xi)\tilde{F}(y|\xi)d\xi$$

$$= \int \rho(\xi)F(x|\xi)d\xi + \int \xi\rho(\xi)F(x|\xi)d\xi$$

$$+ \int \rho(\xi)\tilde{F}(y|\xi)d\xi + \int \xi\rho(\xi)\tilde{F}(y|\xi)d\xi$$

$$= ||\rho||.$$
Inserting the last equality into (20), we have \( |T(W)(\rho) - T(\tilde{W})(\rho)| \leq \alpha||W - \tilde{W}||_G||\rho|| \). Dividing both sides by \( ||\rho|| \), we arrive at the next inequality, which proves the contraction property.

\[
\|T(W) - T(\tilde{W})\|_G = \sup_{\rho \in \mathcal{H}} \frac{|T(W)(\rho) - T(\tilde{W})(\rho)|}{||\rho||} \leq \alpha||W - \tilde{W}||_G.
\]

Since \( T \) is a contraction mapping, the iteration \( W^n(\cdot) = 0 \) and \( W^{n+1}(\rho) = T(W^n)(\rho) \) converges to the solution of the DP equation (13). Moreover, this solution is unique.

ii) Existence of Optimal Solution: We first require two claims.

**Claim 1:** The solution \( W(\cdot) \) of the DP equation (13) is lower semicontinuous over \( \mathcal{H}^+ \).

Since the iteration \( W^n(\cdot) = 0 \) and \( W^{n+1}(\rho) = T(W^n)(\rho) \) converges to the solution of the DP equation (13), it suffices to argue that the sequence \( W^n \) is increasing and each \( W^n \) is continuous. The increasing property follows easily from inspecting the operator \( T \). The continuity property can be proved by induction on \( n \). We suppose that \( W^n(\cdot) \) is continuous and establish the same for \( W^{n+1}(\cdot) \). Note that \( W^0(\cdot) \) is trivially continuous.

For \( \rho_n \rightarrow \bar{\rho} \), consider

\[
|W^{n+1}(\rho_n) - W^{n+1}(\bar{\rho})| = |T(W^n)(\rho_n) - T(W^n)(\bar{\rho})| \\
\leq \int \int L(x,y)f(x|\theta)\rho_n(\theta) - \bar{\rho}(\theta)|d\theta dx \\
+ \alpha \int_0^y |W^n(\rho_n(\cdot)f(x|\cdot)) - W^n(\bar{\rho}(\cdot)f(x|\cdot))| dx \\
+ \alpha |W^n(\rho_n(\cdot)\tilde{F}(y|\cdot)) - W^n(\bar{\rho}(\cdot)\tilde{F}(y|\cdot))|.
\]

Since \( \rho_n(\cdot)f(x|\cdot) \rightarrow \bar{\rho}(\cdot)f(x|\cdot) \), \( \rho_n(\cdot)F(y|\cdot) \rightarrow \bar{\rho}(\cdot)F(y|\cdot) \) and \( W^n \) is continuous, the right-hand side of (21) can be made arbitrarily small as \( \rho_n \rightarrow \bar{\rho} \). That is \( W^{n+1} \) is also continuous. This completes the inductive argument for the continuity of \( W^n \).

**Claim 2:** An order quantity cannot be optimal unless it is less than or equal to

\[
\bar{y} := \frac{bM}{(1 - \alpha)(c - h)} \frac{||\rho||}{\int \rho(\theta)d\theta},
\]

To prove this claim, we show that the solution \( W(\cdot) \) of the DP equation (13) satisfies \( ||W||_G \leq bM/(1 - \alpha) \). We first observe that

\[
(c - h)y \leq L(x,y) \leq \begin{cases} 
\text{cy} & \text{if } x \leq y \\
\text{bx} & \text{if } y \leq x
\end{cases}.
\]

Since \( y = 0 \) is a feasible solution for (13) and \( L(x,0) \leq bx \) by (22), we obtain

\[
W(\rho) \leq \int \int bx f(x|\theta)\rho(\theta)d\theta dx + \alpha W(\rho(\cdot)) \leq b \int \theta \rho(\theta)d\theta + \alpha ||W||_G||\rho|| \\
\leq bM||\rho|| + \alpha ||W||_G||\rho||,
\]

where the second inequality follows from the finite mean demand assumption. Dividing both sides of (23) by \( ||\rho|| \) and taking the supremum over \( \rho \), we obtain \( ||W||_G \leq bM/(1 - \alpha) \).
We insert \( ||W||_B \leq bM/(1 - \alpha) \) into (23) to obtain
\[
W(\rho) \leq bM ||\rho|| \left( 1 + \frac{\alpha}{1 - \alpha} \right) = \frac{bM ||\rho||}{1 - \alpha}.
\] (24)

On the other hand, a simple lower bound on the cost of ordering \( y \) can be deduced by using (22) and ignoring the future costs as
\[
\int \int (c - h) y f(x|\theta) \rho(\theta) d\theta dx = (c - h) y \int \rho(\theta) d\theta \leq W(\rho).
\] (25)

Combining (24) and (25), we arrive at
\[
(c - h) y \int \rho(\theta) d\theta \leq \frac{bM ||\rho||}{1 - \alpha},
\]
which establishes Claim 2.

As a consequence of the lower semicontinuity of \( W \) in Claim 1, we obtain the lower semicontinuity of the map \( y \mapsto T_y(W)(\rho) \). Moreover, the optimal order quantity belongs to the interval \([0, \tilde{y}]\) by Claim 2. Thus, \( \inf_y T_y(W)(\rho) \) must be attained by an optimal \( y^* \in [0, \tilde{y}] \). \( \square \)

**Proof of Theorem 4.2**: To obtain \( y^*(\rho) \geq y^M(\rho) \) for each arbitrary \( \rho \), it suffices to show
\[
\frac{d}{dy} \left\{ \int \int L(x, y) f(x|\theta) \rho(\theta) d\theta dx + \alpha \int_0^y W(\rho(\cdot)f(x|\cdot)) dx + \alpha W(\rho(\cdot)\bar{F}(y|\cdot)) \right\} \leq \frac{d}{dy} \left\{ \int \int L(x, y) f(x|\theta) \rho(\theta) d\theta dx \right\}.
\]

After taking the derivative of the integral, we arrive at an equivalent inequality
\[
\frac{d}{dy} W(\bar{F}(y|\cdot)\rho(\cdot)) \leq -W(f(y|\cdot)\rho(\cdot)). \tag{26}
\]

In this proof, we write \( y^*(\rho; W) \) when we want to make explicit the value function appearing on the right-hand side of (14).

We consider the iteration \( W^0(\rho) = 0 \) and \( W^{n+1}(\rho) = T(W^n)(\rho) \). Since \( T \) is a contraction mapping, \( W^n \) converges to the solution of (13). Therefore we can state and inductively prove (26) in terms of \( W^{n+1} \)
\[
W^{n+1}(f(y|\cdot)\rho(\cdot)) \leq -\frac{d}{dy} W^{n+1}(\bar{F}(y|\cdot)\rho(\cdot)), \tag{27}
\]
knowing that (27) holds trivially for \( W^0(\rho) = 0 \).

The key observation here is that we can insert a third expression between the left-hand side and the right-hand side of (27). This expression is
\[
\left\{ \int \int L(x, \eta) f(x|\theta) f(y|\theta) \rho(\theta) d\theta dx + \alpha \int_0^\eta W^n(f(y|\cdot)\rho(\cdot)f(x|\cdot)) dx + \alpha W^n(f(y|\cdot)\rho(\cdot)\bar{F}(\eta|\cdot)) \right\}.
\]
This expression is minimized at \( \eta = y^*(f(y|\cdot)\rho(\cdot); W^n) \) and becomes equal to \( W^{n+1}(f(y|\cdot)\rho(\cdot)) \). When evaluated at \( \eta = y^*(\bar{F}(y|\cdot)\rho(\cdot); W^n) \), it is larger than or equal to \( W^{n+1}(f(y|\cdot)\rho(\cdot)) \), which is the left-hand side of (27). To show that this expression at \( \eta = y^*(\bar{F}(y|\cdot)\rho(\cdot); W^n) \) is less than or equal to the right-hand side of (27), we can take the derivative of \( W^{n+1}(\bar{F}(y|\cdot)\rho(\cdot)) \) and use the envelope theorem.
(see, e.g., Derzko, Sethi and Thompson [7]). Finally, the induction hypothesis, that is (27) with index $n$, yields the desired inequality. More details on these steps can be found in Bensoussan et al. [6]. These steps complete the induction argument as well as the proof. □

Acknowledgements. We would like to thank the anonymous referees very much for their valuable comments and suggestions.

REFERENCES

[1] A. Bensoussan, Stochastic control of partially observable systems, Cambridge University Press, 1992.
[2] (MR2168741) A. Bensoussan, M. Çakanyıldırım and S.P. Sethi, On the optimal control of partially observed inventory systems, Comptes Rendus Mathematique, 341, (2005), 419–426.
[3] (MR2324430) A. Bensoussan, M. Çakanyıldırım and S.P. Sethi, A multiperiod newsvendor problem with partially observed demand, Mathematics of Operations Research, 32, (2007), 322-344.
[4] (MR2466408) A. Bensoussan, M. Çakanyıldırım and S.P. Sethi, Partially observed inventory systems: The case of zero balance walk, SIAM Journal of Control and Optimization, 46, (2007), 176-209.
[5] A. Bensoussan, M. Çakanyıldırım, A. Royal and S.P. Sethi, Bayesian and Adaptive Controls for a Newsvendor Facing Exponential Demand, To appear in Risk and Decision Analysis.
[6] A. Bensoussan, M. Çakanyıldırım and S.P. Sethi, A Note on “the censored newsvendor and the optimal acquisition of information”, To appear in Operations Research.
[7] (MR0745788) N.A. Derzko, S.P. Sethi and G.L. Thompson, Necessary and Sufficient Conditions for Optimal Control of Quasilinear Partial Differential Systems, Journal of Optimization Theory and Applications, 43, (1984), 89-101.
[8] (MR1910287) X. Ding, M.L. Puterman and A. Bisi, The censored newsvendor and the optimal acquisition of information, Operations Research, 50, (2002), 517-527
[9] M.A. Lariviere and E.L. Porteus, Stalking information: Bayesian inventory management with unobserved lost sales Management Science, 45, (1999), 346-363.
[10] (MR2193876) X. Lu, J.S. Song and K. Zhu, On “the censored newsvendor and the optimal acquisition of information”, Operations Research, 53, (2005), 1024-1027.
[11] (MR2455712) X. Lu, J.S. Song and K. Zhu, Analysis of Perishable inventory systems with censored demand data, Operations Research, 56, (2008), 1034-1038.
[12] (MR0649094) G.E. Monahan, A survey of partially observable Markov decision processes, Management Science, 28, (1982), 1-16.
[13] M. Zakai, On the optimal filtering of diffusion processes, Z. Wahrsch. verw. Gebiete, 11, (1968), 230-243.

E-mail address: alain.bensoussan@utdallas.edu
E-mail address: metin@utdallas.edu
E-mail address: sethi@utdallas.edu