Inventory Control Driven by Demand Data: Optimality and Computation of Base Stocks

Alain Bensoussan
International Center for Decision and Risk Analysis (ICDRiA), Naveen Jindal School of Management, University of Texas at Dallas, Richardson, TX 75080-5298, USA, alain.bensoussan@utdallas.edu
& Department of Systems Engineering and Engineering Management, the City University of Hong Kong

Suresh P. Sethi
Naveen Jindal School of Management, University of Texas at Dallas, Richardson, Texas 75080-3021, USA, sethi@utdallas.edu

Abdoulaye Thiam
Department of Mathematical Sciences, Natural Sciences & Mathematics, & (ICDRiA), University of Texas at Dallas, Richardson, TX 75080-3021, USA, abdoulaye.thiam@utdallas.edu

Janos Turi
Department of Mathematical Sciences, Natural Sciences & Mathematics, University of Texas at Dallas, Richardson, TX 75080-3021, USA, janos.turi@utdallas.edu

This paper advances significantly the literature on the optimality of the base stock policy by generalizing the demand distribution and beginning with a completely general belief prior to be updated as demands are observed over time. As the value function depends on the belief, the functional Bellman equation is infinite-dimensional. We use unnormalized probabilities to linearize it and derive a functional equation for the derivative of the value function. This provides a constructive approach to obtain the base stock as well as the value function. We completely characterize the way the base stock depends on the belief, and implement it in two important cases. In the case of conjugate probabilities, we show rigorously that the infinite-dimensional problem reduces to one in terms of a finite-dimensional sufficient statistic. We solve numerically an example of Weibull demand. The second case considers the demand to come from one of two possible distributions, but we don't know which. This gives a functional equation in two hyperparameters. We develop two approximation schemes to solve it, obtain the formulas for the base stock, and show numerically that both procedures converge and provide nearly the same base stocks. Finally, our methodology can be used for examples such as the case of multiple possible demands and the case when a fixed ordering cost is present.

Key words: stochastic dynamic programming, unnormalized probabilities, demand learning, base stock policy, random demands.

1. Introduction

The optimality of the base stock policy when the demand distribution is known is among the most fundamental results in the inventory theory. However, the assumption of known demand conflicts with many real situations when some degree of uncertainty exists in the mind of the decision maker about the distribution of demand. In such situations, a Bayesian framework has been used since intuition and previous experience which represent the initial state of information can be expressed by a prior distribution, and Bayes' rule provides a way to incorporate new information as it becomes available into the decision model.

In this paper, we advance significantly the literature on the optimality of the base stock policy with demand learning by extending it to cases with general demand distributions and beginning with a completely general belief prior to be updated as the demand history unfolds over time. We consider a standard discrete-time infinite-horizon inventory problem, but with a general demand that is not known, and allow unmet demands to be backlogged so that the realized demands are fully observed (Section 4). We assume the demand to depend on an unknown (scalar or a vector) parameter and an initial belief given in terms of a general probability density function, as a point of departure from the extant literature, reviewed in Section 2 and EC.1. This means that the state equation evolves in the infinite-dimensional state space of the current inventory and the current belief density, which get updated after the ordering decision and the demand realization in the period. As we know that in the standard inventory problem when there is no unknown parameter, there is no learning, and the optimal feedback is a base stock policy (Section 3). That cannot work anymore since a feedback rule based on the current inventory alone would entirely ignore the learning process. Thus, the goal here is to obtain a similar result when there is learning, with a base stock now depending on the current belief probability updated from observing the past demands. We accomplish this main finding of the paper by analyzing the functional Bellman equation resulting from the dynamic programming formulation of the infinite-dimensional inventory problem under consideration. Moreover, the value function is now a functional that not only depends on the usual current inventory level, but also on the current infinite-dimensional belief, and we use a monotonicity argument in Section 6 to show that it is the unique solution of the Bellman equation in an appropriate Banach space.

The ensuing analysis represents another important methodological contribution of the paper that can be applied in other similar contexts. To facilitate the study of the Bellman equation, we extend the domain of belief functions to positive functions, called unnormalized probabilities, which are not necessarily probability densities. This extension results in a simpler, equivalent Bellman equation in which belief updating becomes linear. Since the Bellman equation involves an inf operation to obtain the optimal solution, we need to differentiate the value function. Instead of this usual approach, we derive a functional equation for the derivative of the value function and then obtain it as the fix point of the equation directly by an iterative approach. This procedure yields not only the derivative of the value function, but also the base stock. Moreover, the value function can be recovered from these as well.

Our next contribution is in applying the developed theory to two special cases where we assume the belief function to depend on a finite set of hyperparameters, thus reducing the infinite-dimensional problem to a finite-dimensional one, the dimension being the number of the hyperparameters. The first case is to choose the belief function to be a conjugate prior of the demand density, and show that the belief can now be expressed in terms of its hyperparameters which can be updated based on the observed demands, and thus the base stock becomes a function of this sufficient statistic. We illustrate these results by assuming a Weibull demand whose scale parameter is not known and whose conjugate prior is the family of Gamma densities characterized by two hyperparameters. We characterize the base stock in terms of these two hyperparameters and obtain them numerically. We obtain the optimal orders in an example for, say, the first five periods. We also validate the model by numerical simulation. That is, we create demands according to an assumed true Weibull density, start with an initial belief about the hyperparameters, update them based on the simulated demands, obtain the sequence of the optimal base stocks in terms of the updated hyperparameters, and then show that the sequence converges to the true base stock.

Our second special case considers the demand to come from one of two possible distributions, but we don't know which. This gives rise to a single hyperparameter that expresses the ratio of the weights assigned to the two distributions. We use our theory to write the functional equation satisfied by the derivative of the value function in this case, develop two approximation schemes to solve it, and thereby obtain the base stock. We apply these schemes to a particular case of high or low exponential demand distributions and obtain the formulas for the base stock. We show numerically that both procedures converge and provide nearly the same base stocks.

The plan of this paper is as follows. In Section 2, we review the related literature and delineate our contributions to the literature. Section 3 recalls the classical base stock policy results in inventory control problems with backlog allowed. In Section 4, we formulate an inventory problem with a general demand depending on a parameter, introduce a general belief density of the parameter, and use the Bayesian learning process to update the density based on the demands observed over time. In Section 5, we use dynamic programming to obtain the functional equation and obtain bounds on its solution to ensure a unique solution. In Section 6, we prove that the value function is the only solution of the Bellman equation and that there exists an optimal feedback policy. In Section 7, we show that the optimal feedback policy can be expressed as a base stock policy where the base stock depends only on the current belief function. In Section 8, we treat the special case of conjugate probabilities with the belief modeled by the conjugate prior of the demand distribution, and characterize the dependence of the base stock on the set of hyperparameters of the belief density. We apply the theory to a particular case of the Weibull-Gamma conjugates. In Section 9, we consider the case when demand comes from one of two possible distributions, but we don't know which. We develop two approximation schemes to obtain the base stock and show numerically that both procedures converge to nearly the same base stocks. Section 10 concludes the paper. An e-companion contains the proofs of results, derivation of some of the equations, and a review of the additional literature that is related but only tangentially to the specific topic of the base stock policy.

2. Review of Related Literature

Inventory problems with Bayesian learning of an unknown demand have their origin in the classical papers of Dvoretzky et al. (1952), Scarf (1959) and Scarf (1960a), with Scarf's work being the seminal as well as the most related to our paper. Scarf (1959) pioneered the Bayesian approach to show that the base stock policy remains optimal in the presence of demand learning, under the assumption that the prior distribution is in the conjugate family to the form of the unknown demand distribution. He considers an infinite-horizon inventory problem with demand densities to be in the exponential class $g(z|\xi) = \beta(\xi)r(z)e^{-\xi z}$, where $r(z)$ is bounded and strictly positive for $z > 0$. Also, $r(z) = 0$ for $z < 0$, since demands are nonnegative. It is then known that all of the relevant information in any given period may be summarized in a single sufficient statistic, namely the mean of the past demand observations as the only hyperparameter. He sets up the functional Bellman equation for the value function depending on the current inventory level and the current update of the hyperparameter (two variables), approximates it by a recursion over a finite horizon N, and obtains the base stock policy for the approximate problem. He proves that the approximate base stock obtained for any given period n converges to the base stock in period n of the original infinite horizon problem as N goes to infinity, which in turn converges to the true base stock as n approaches infinity. He also provides an asymptotic expansion of the base stock when the number of demand observations is large, and shows that for large n the base stock becomes close to the true base stock and the error term is of the order of $1/n$. Karlin (1960) and Iglehart (1964) extend Scarf's results to the range family of densities. Owing to the monotone likelihood ratio property satisfied by these distributions, it is also shown that the base stock is nondecreasing in their respective sufficient statistic.

Scarf (1960a) notes that the functions of two variables arising in Scarf (1959) are difficult to compute recursively. So he makes a simplifying assumption $\beta(\xi)r(z) = \xi^m z^{m-1}$, which makes the demand density a Gamma density. This enables him to determine the optimal base stock levels by the recursive computation of functions of only the inventory level (one variable), and thus without any approximation. Azoury (1985) extends the reduction result in a natural way to demand densities such as Gamma, uniform, and Weibull. Lovejoy (1990) accomplishes a further reduction of the dynamic program with a single state variable to one with a "zero dimensional" state space, i.e., a static optimization problem. This is accomplished by imposing assumptions in addition to those of Scarf (1960a) and Azoury (1985) that provide for a myopic optimal policy, which is given by a simple critical fractile stock level as the base stock in each period.

Building on the works of Subrahmanyan and Shoemaker (1996) and Petruzzi and Dada (2002), Zhang and Chen (2006) allow for price as an additional decision variable. They assume that the demand in any given period is a basic random demand minus a deterministic component that increases linearly in price. They follow Azoury (1985) to reduce the dimension of the state space of the dynamic programming recursion. They show that a base stock list price policy, coined in the Porteus (1990) inventory theory survey, is optimal. Other related papers on dynamic pricing with demand learning are Aviv and Pazgal (2005), Bisi and Dada (2007), Farias and Van Roy (2010) and Araman and Caldentey (2009).

Larson et al. (2001) treat both finite and infinite horizon problems by considering a nonparametric Bayesian approach in which a firm's prior information about the demand distribution is characterized by a Dirichlet process prior on the space of distributions. We note that this setting makes the hyperparameter also a probability measure. Given such a Dirichlet process prior, any distribution whose support is included in the support of the measure characterizing the prior can be approximated as a posterior demand distribution, under the topology of weak convergence. Thus, the approach provides a consistent model of learning provided that the support of the true demand distribution is in a subset of the support of the prior. As the authors allow for fixed ordering cost, their focus is on showing the optimality of a history-dependent (s, S) policy, which is known to reduce to a base stock policy when the fixed cost is zero. Furthermore, they show that if the optimal history-dependent (s, S) policies take a limit as demand information accumulates, then they converge to the (s, S) policy that is optimal for the true underlying demand distribution. The authors note that their approach does not smooth beliefs as in the conjugate family settings where observing a high outcome typically implies that other high outcomes are also more likely. In this regard, the approach is likely to be most useful when observations occur frequently, or when the absence of smoothing is of minor concern. Indeed, their model allows demands to be observed more frequently than once every period.

As we can see that previous relevant research, with the exception of Larson et al. (2001), specify the information about demand typically by some conjugate prior on the unknown demand parameters and updates via Bayes' rule. Typically, the choice of a conjugate family of distributions places restrictions on the prior beliefs that can be accommodated and the true demand distributions that can be allowed. For example, it is difficult under this setting to allow for bimodal priors or bimodal true demand densities. Further conditions imposed to achieve dimensionality reduction also restricts the choice of distributions or the priors. On the other hand, given a Dirichlet process prior as in Larson et al. (2001), any demand distribution whose support is included in the support of the measure characterizing the prior can be approximated as a posterior, but if we are given an arbitrary belief, there may not be a Dirichlet process prior that can closely approximate it. In particular, if a demand distribution is a Dirichlet process, then with probability one it is discrete. This has the limitation that if the set of possible demand distributions is a subset of the continuous distributions, then the Dirichlet process assigns probability zero to the true set. In other words, while the Dirichlet process is rather flexible as a model of the underlying demand distribution, it is less so as a model of beliefs. Furthermore, as Larson et al. (2001) note that their approach is likely to be most useful when observations occur frequently. If that is the case, the approach may not be suitable in common situations where the demand realizes once every period. Moreover, the hyperparameter in the Dirichlet approach being itself a probability measure, the approach may not be convenient for computational purposes.

We can now articulate our contributions in relation to the literature. We advance the research on the optimality of the base stock policy in the presence of demand learning significantly by assuming a general demand depending on an unknown (scalar or a vector) parameter and an initial belief given in terms of a general probability density function. In Section 8, we specialize our theory to the case of conjugate families of distributions, and present results that generalize all previous research that has assumed the firm's information about demand to be specified by some conjugate prior on the unknown parameters and updated via Bayes' rule. While we develop the main theory with a parametric demand, our method can handle nonparametric demands since our belief can be completely general. That allows us to consider a finite set of possible nonparametric demand distributions containing the true distribution and a belief modeled by a Dirac mass associated with each distribution along with its coefficient representing the weight assigned to that distribution. Then the coefficients are the hyperparameters that can be updated as demands get observed. Indeed, we illustrate this setup in Section 9 with two arbitrary demand distributions as a special case of our general theory.

Thus, in connection with proving that a base stock policy is optimal in the presence of demand learning, our paper is arguably the most general one to date. In addition, we also make major methodological contributions which we have already summarized in the previous section. Before we conclude this section, we would like to mention that there is a considerable literature on partially observed inventory control problems with demand learning that is not devoted to extending the optimality of base stock, base stock list price, and (s,S) policies. We review very briefly this literature in EC.1.

3. Classical Base Stock Policy

We recall the standard discrete-time inventory control problem over an infinite horizon. The inventory manager IM decides on his inventory order v_n, $n=1,2,\cdots$, at the beginning of period n that is delivered in the period before the demand in the period materializes. The demands D_n, $n=1,2,\cdots$, are independent and identically distributed random variables with the probability density function (PDF) on $\mathbb{R}^+$ denoted by $g(z)$. We define $G(z)=\int_0^z g(t)dt$, the cumulative distribution function (CDF), and $\bar{G}(z)=1-G(z)$. In this section, we assume this distribution to be known, an assumption that will be relaxed in the remainder of the paper.

The demand is satisfied to the extent possible from the inventory on hand in each period, the excess inventory or the backlog, as the case may be, is carried over to the next period. Then with x_n denoting the inventory at the beginning of period n, the inventory evolves according to

$$\left| \begin{array}{l} x_{n+1}=x_n+v_n-D_n, \\ x_1=x. \end{array} \right. \tag{1}$$

The order $v_n=v(x_n)$ is a function of the inventory at the beginning of period n. It is called a feedback control in the engineering literature. It is well known that it is not a restriction to limit order decisions to feedback controls. Denoting by $V=\{v_1,\cdots,v_n,\cdots\}$, the feedback to be chosen, the payoff to minimize is

$$J_x(V)=E\sum_{n=1}^{\infty}\alpha^{n-1}(hx_n^+ + px_n^- + cv_n). \tag{2}$$

The parameters p, h and c are, respectively, the penalty cost of backlog per unit per period, the inventory holding cost per unit per period, and the unit ordering cost. The parameter α is the discount factor. We define the value function

$$\Phi(x) = \inf_{V} J_x(V). \tag{3}$$

Then, $\Phi(x)$ is solution of the Bellman equation

$$\Phi(x) = hx^+ + px^- + \inf_{v>0} \left[cv + \alpha E\Phi(x+v-D)\right]. \tag{4}$$

We make a standard assumption

$$\frac{\alpha p}{1-\alpha} > c \tag{5}$$

to rule out the trivial solution of no ordering and backlogging all demands to be optimal. This is easy to see by knowing that the right hand side is the cost of ordering a unit of inventory in the current period and the left hand side is the present value of a unit backlog carried out from the next period on all the way to infinity. In other words, if (5) does not hold, then the base stock, as it were, will be $-\infty$.

The solution is obtained as follows. Let S, called the base stock, to be the unique solution of

$$\bar{G}(S) = \frac{c - \alpha c + \alpha h}{\alpha(h+p)}. \tag{6}$$

The optimal feedback is then

$$\hat{v}(x) = \left| \begin{array}{l} S - x, \;\text{ if }\; x < S, \\ \\ 0, \;\text{ if }\; x \geq S. \end{array} \right. \tag{7}$$

For $x \leq S$,

$$\Phi(x) = hx^+ + px^- - cx + H, \tag{8}$$

with

$$H = \frac{cS(1-\alpha) + c\alpha\bar{D} + hE(S-D)^{+} + pE(S-D)^{-}}{1-\alpha}. \tag{9}$$

For $x \geq S$, $\Phi(x)$ is solution of the linear integral equation

$$\Phi(x) = hx^{+} + px^{-} + \alpha \int_{0}^{+\infty} \Phi(x-t)g(t)dt. \tag{10}$$

Note that from the definition of H, the function $\Phi(x)$ is continuous at S. Define

$$u(x) = \Phi'(x) - h\mathbb{1}_{x>0} + p\mathbb{1}_{x<0} + c. \tag{11}$$

Then $u(x) = 0$ for $x \leq S$, and for $x \geq S$ we can see from (10) that it satisfies the relation

$$u(x) = c(1-\alpha) + \alpha h - \alpha(h+p)\bar{G}(x) + \alpha \int_{0}^{+\infty} u(x-t)g(t)dt. \tag{12}$$

By use of the contraction mapping theorem, equation (12) defines a unique positive bounded function on $[S, +\infty)$ with the property that $u(S) = 0$. We can then extend $u(x)$ by setting it as 0 for $x \leq S$. This defines a nonnegative function $u(x)$ which is continuous and bounded on $\mathbb{R}$. Then the function $\Phi(x)$ of (10) for $x \geq S$ is given by the relation

$$\Phi(x) = \Phi(S) - (c-h)(x-S) + \int_{S}^{x} u(t)dt, \quad x \geq S. \tag{13}$$

We have thus defined a function $\Phi(x)$ by (8) and (13) which is $\mathcal{C}^1$ and it solves the functional Bellman equation (4). See (Bensoussan 2011) for details.

4. Bayesian Learning

We now assume that the IM does not have full knowledge of the probability density $g(z)$ of the demand. More precisely, we write $g(z|\xi)$ with ξ a parameter not known to the IM, but he will learn about it by observing the demands over time. We can no longer, as in Section 3, work with a feedback rule based on the current inventory as that would ignore the learning process. Since the information comes from observing the demands, we introduce the filtration

$\mathcal{D}^n = \sigma(D_1, \cdots, D_n),\ \ n \geq 1$. The order quantity v_n at time n is $\mathcal{D}^{n-1}$ measurable. At the initial time 1, there is no information except that the inventory $x_1 = x$, which is given. However, the IM has an initial belief on ξ, which we express by the PDF $f(\xi)$ on $\mathbb{R}$. This belief will evolve through the learning process. From the inventory evolution

$$\left| \begin{aligned} & x_{n+1} = x_n + v_n - D_n, \\ & x_1 = x, \end{aligned} \right. \tag{14}$$

we can see that the inventory process x_n is also $\mathcal{D}^{n-1}$ measurable.

The belief density is updated by the Bayesian learning process. For this, let us define the updated belief after observing the demands $D_1, \cdots, D_{n-1}$ as $f_n(\xi)$ with $f_1(\xi) = f(\xi)$. Then we can write

$$f_n(\xi) = \frac{f(\xi) g(D_1|\xi) \cdots g(D_{n-1}|\xi)}{\int f(\eta) g(D_1|\eta) \cdots g(D_{n-1}|\eta) d\eta} \tag{15}$$

by observing that the numerator gives the joint probability density of $\xi, D_1, \cdots, D_{n-1}$, since the demands are independent given the parameter ξ, and the denominator is the joint probability density of the demands as they are no longer independent in the presence of the parameter. Using $f_n(\xi)$, it is easy to obtain the conditional density of D_n, given $D_1, \cdots, D_{n-1}$, as

$$g(D_n|\mathcal{D}^{n-1}) = \int f_n(\eta) g(D_n|\eta) d\eta. \tag{16}$$

We can now use Bayes' theorem to obtain

$$f_{n+1}(\xi) = \frac{g(D_n|\xi) f_n(\xi)}{g(D_n|\mathcal{D}^{n-1})} = \frac{g(D_n|\xi) f_n(\xi)}{\int g(D_n|\eta) f_n(\eta) d\eta}. \tag{17}$$

The pair $\{x_n, f_n(.)\}$ is a stochastic process adapted to $\mathcal{D}^{n-1}$. Setting $y_n = x_n + v_n$, we can write (14) and (17) as

$$x_{n+1} = y_n - D_n, \quad x_1 = x, \tag{18}$$

$$f_{n+1}(\xi) = f_n(\xi) \frac{g(D_n|\xi)}{\int g(D_n|\eta) f_n(\eta) d\eta}, \quad f_1(\xi) = f(\xi). \tag{19}$$

The pair $(x_n,\, f_n(.))$ becomes the state of our dynamic system with the state space $\mathbb{R}\times\mathcal{L}(\mathbb{R})$, where $\mathcal{L}(\mathbb{R})$ is the set of probability densities on $\mathbb{R}$. With this new system state, we can now define the inventory control problem of IM. We set the decision $V=\{y_1,\cdots,y_n,\cdots\}$ with $y_1=y\geq x$, $y_n\geq x_n$, where y_n-x_n denotes the order quantity in period n, and we define the payoff

$$J_{x,f(.)}(V)=\sum_{n=1}^{+\infty}\alpha^{n-1}E\big[hx_n^{+}+px_n^{-}-cx_n+cy_n\big]. \tag{20}$$

The value function can be denoted as

$$\Phi(x,f(.))=\inf_{V}J_{x,f(.)}(V), \tag{21}$$

where $(x,f(.))$ is the initial state.

5. Dynamic Programming and Control

In this section we use dynamic programming and write the functional Bellman equation satisfied by the value function. As the Bellman equation may have many solutions, we develop conditions under which the value function becomes its unique solution. In Subsection 5.1, we obtain lower and upper bounds on the value function, which will allow us in Section 6 to show that the solution of the Bellman equation is unique when restricted to these bounds, and that solution is indeed the value function. Furthermore, we can use these bounds also to obtain an upper bound on the order-up-to level decision y. Note that the beginning inventory level x in is naturally a lower bound on the decision y. These decision bounds allow us to obtain a modified Bellman equation. In subsection 5.2, we obtain an intermediate comparison result that would help us in proving the main result of Section 6.

The functional $\Phi(x,f(.))$ defined in (21) satisfies the Bellman equation

$$\Phi(x,f(.))=hx^{+}+px^{-}-cx+\inf_{y\geq x}\Big[cy+\alpha E\Phi\Big(y-D,\frac{f(.)g(D|.)}{\int f(\eta)g(D|\eta)d\eta}\Big)\Big], \tag{22}$$

and the unique solution of it will be the value function as defined in (21).

5.1. Bounds on the Value Function

In order to have the value function (21) to be the only solution of (22), we need some conditions which in our case appear by imposing upper and lower bounds on that unique solution. In other words, we would like to have the Bellman equation to take a unique solution when restricted to these bounds. It turns out that we can find these bounds directly from the definition of the value function, i.e., from (20)-(22), as shown in EC.2. These bounds are presented in the following lemma.

LEMMA 1. *The value function $\Phi(x, f(.))$ defined in (21) satisfies*

$$\Big(\frac{hx^+}{1-\alpha}+px^- - \frac{\alpha h}{(1-\alpha)^2}\int \varphi(\eta)f(\eta)d\eta\Big)^+ \le \Phi(x,f(.)) \le \frac{hx^+}{1-\alpha}+(p+c)x^- + \frac{\alpha(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta, \tag{23}$$

where we assume that $\int \varphi(\eta)f(\eta)d\eta < +\infty$, *with*

$$\varphi(\eta)=\int_0^{+\infty} \bar{G}(z|\eta)dz. \tag{24}$$

Next we show that if a solution of (22) is to be the value function, then the bounds in (23) imply an upper bound on the decision $y \ge x$. For a y that achieves the inf in (22), the right hand side with that y equals $\Phi(x, f(.))$. Then we can use the upper bound in (23) to obtain

$$hx^+ + px^- - cx + cy + \alpha E\Phi\Big(y-D, \frac{f(.)g(D|.)}{\int f(\eta)g(D|\eta)d\eta}\Big) \le \frac{hx^+}{1-\alpha}+(p+c)x^- + \frac{\alpha(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta.$$

By noting that $x + x^- = x^+$, we can rewrite it as

$$cy + \alpha E\Phi\Big(y-D, \frac{f(.)g(D|.)}{\int f(\eta)g(D|\eta)d\eta}\Big) \le cx^+ + \frac{\alpha hx^+}{1-\alpha} + \frac{\alpha(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta. \tag{25}$$

On the other hand, if we put $y = x$ in (22), then its right hand side will be larger than or equal to $\Phi(x, f(.))$, and thus also larger than or equal to the lower bound in (23). Therefore, we have

$$\begin{aligned}\frac{hx^+}{1-\alpha}+px^- - \frac{h\alpha}{(1-\alpha)^2}\int \varphi(\eta)f(\eta)d\eta &\le \Big(\frac{hx^+}{1-\alpha}+px^- - \frac{h\alpha}{(1-\alpha)^2}\int \varphi(\eta)f(\eta)d\eta\Big)^+ \\ &\le hx^+ + px^- + \alpha E\Phi\Big(y-D, \frac{f(.)g(D|.)}{\int f(\eta)g(D|\eta)d\eta}\Big).\end{aligned}$$

So,

$$-\alpha E\Phi\Big(y-D, \frac{f(.)g(D|.)}{\int f(\eta)g(D|\eta)d\eta}\Big) \le -\frac{\alpha hx^+}{1-\alpha} + \frac{\alpha h}{(1-\alpha)^2}\int \varphi(\eta)f(\eta)d\eta. \tag{26}$$

By adding (25) and (26) and dividing by c, we see that the decision y can be chosen to satisfy

$$y \leq x^+ + \frac{1}{c}\frac{\alpha}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\int \varphi(\eta)f(\eta)d\eta. \tag{27}$$

Hence, we can define the following *admissibility interval* for y :

$$y \in \mathcal{I}_{x,f(.)} \stackrel{\text{def}}{=} \Big[x, x^+ + \frac{1}{c}\frac{\alpha}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\int \varphi(\eta)f(\eta)d\eta\Big]. \tag{28}$$

We can then replace the domain of $y \geq x$ in (22) by the domain (28) to obtain the modified Bellman equation

$$\Phi(x,f(.)) = hx^+ + px^- - cx + \inf_{y\in\mathcal{I}_{x,f(.)}}\Big[cy + \alpha\int_0^{+\infty}\Phi\Big(y-z, \frac{f(.)g(z|.)}{\int f(\eta)g(z|\eta)d\eta}\Big)\int f(\eta)g(z|\eta)d\eta dz\Big]. \tag{29}$$

In Subsection (6.2), we will show that the unique solution of (23) and (29) is the value function.

5.2. A Comparison Result

We provide here an important comparison result between the solutions of (29) satisfying the bounds in (23) and the value function $\inf_V J_{x,f(.)}(V)$ defined in (21). To avoid confusion, we temporarily call the value function by

$$\Phi^*(x,f(.)) = \inf_V J_{x,f(.)}(V).$$

Proposition 1. *Any solution $\Phi(x,f(.))$ of the functional equation (29) such that (23) holds satisfies*

$$\Phi(x,f(.)) \leq \Phi^*(x,f(.)) = \inf_V J_{x,f(.)}(V). \tag{30}$$

Its proof is relegated to EC.3. This result will be required to establish our main result of the next section.

6. Main Result

In this section, we prove that the value function is the unique solution of the Bellman equation and there exists an optimal feedback policy. For proving these results stated as Theorem 1, we make the following assumptions:

$$g(z|\eta) \le C(z), \, \forall z > 0, \, \forall \eta, \tag{31}$$

$$C(z) \text{ continuous}, \tag{32}$$

$$\varphi(\eta) g(z|\eta) \le C, \text{ independent of } z, \eta, \tag{33}$$

where $\varphi(\eta)$ is defined in (24).

THEOREM 1. *Assume (31)-(33). For any probability density $f(.)$ on $\mathbb{R}$ such that $\int \varphi(\eta) f(\eta) d\eta < +\infty$, then the value function $\Phi(x, f(.))$ is the unique functional such that (23) and (29) hold. Moreover, there exists a $\hat{y}(x, f(.))$ that attains the inf in (29).*

For proving Theorem 1, it is convenient to work with unnormalized probability introduced in the next subsection. We will see that the belief updating in (29) becomes linear when expressed in terms of unnormalized probability, which will facilitate a further study of the Bellman equation (29) once it is transformed to (39) in the next subsection.

6.1. Unnormalized Bellman Equation

Consider the space of functions on $\mathbb{R}$ such that

$$||f||_\varphi = \int |f(\eta)| d\eta + \int \varphi(\eta) |f(\eta)| d\eta < +\infty. \tag{34}$$

We denote this space as $L^1_\varphi(\mathbb{R})$, and it is a Banach space for the norm $||f||_\varphi$. The subset of positive functions denoted by $L^1_{\varphi+}(\mathbb{R})$ is closed. We now introduce a simpler but equivalent equation, called the unnormalized Bellman equation. The idea is to extend the domain of the function $f(.)$ to unnormalized probabilities, which are simply positive functions which include probability densities. So a functional on $\mathbb{R}$ and probability densities is extended to a functional on $\mathbb{R} \times L^1_{\varphi+}(\mathbb{R})$ by the formula

$$W(x, f(.)) = \Phi\Big(x, \frac{f(.)}{\int f(\eta) d\eta}\Big) \int f(\eta) d\eta. \tag{35}$$

Note that $W(x,f(.))$ and $\Phi(x,f(.))$ coincide when $f(.)$ is a probability density. From (35), we can see that the decision y satisfies

$$x\int f(\eta)d\eta \leq y\int f(\eta)d\eta \leq \Big(x^+ + \frac{1}{c}\frac{\alpha}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\frac{\int \varphi(\eta)f(\eta)d\eta}{\int f(\eta)d\eta}\Big)\int f(\eta)d\eta.$$

So,

$$y \in \mathcal{L}_{x,f(.)} \stackrel{\text{def}}{=} \Big[x, x^+ + \frac{1}{c}\frac{\alpha}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\frac{\int \varphi(\eta)f(\eta)d\eta}{\int f(\eta)d\eta}\Big]. \tag{36}$$

Also by (35), we see that

$$W(x,\mu f(.)) = \Phi\Big(x, \frac{\mu f(.)}{\int \mu f(\eta)d\eta}\Big)\int \mu f(\eta)d\eta = \mu W(x,f(.)), \ \ \forall \mu > 0, \tag{37}$$

which, in turn, gives

$$\begin{aligned}&\Phi(y-z, \frac{f(.)g(z|.)}{\int f(\eta)g(z|\eta)d\eta})\int f(\eta)g(z|\eta)d\eta\\ &= W(y-z, \frac{f(.)g(z|.)}{\int f(\eta)g(z|\eta)d\eta})\int f(\eta)g(z|\eta)d\eta = W(y-z, f(.)g(z|.)).\end{aligned} \tag{38}$$

Hence, from (29) and (35), we get the unnormalized Bellman equation

$$W(x,f(.)) = (hx^+ + px^- - cx)\int f(\eta)d\eta + \inf_{y\in\mathcal{L}_{x,f(.)}}\big[cy\int f(\eta)d\eta + \alpha\int_0^{+\infty} W(y-z, f(.)g(z|.))dz\big], \tag{39}$$

and by Lemma 1, we can easily obtain the following bounds of $W(x,f(.))$:

$$\begin{aligned}&\Big((\frac{hx^+}{1-\alpha} + px^-)\int f(\eta)d\eta - \frac{h\alpha}{(1-\alpha)^2}\int \varphi(\eta)f(\eta)d\eta\Big)^+\\ &\leq W(x,f(.))\\ &\leq (\frac{hx^+}{1-\alpha} + (p+c)x^-)\int f(\eta)d\eta + \frac{\alpha(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta.\end{aligned} \tag{40}$$

To conclude, it is clear that a solution of (39) and (40) satisfies (37). Also (39) and (40) are equivalent to (29) and (23), respectively, and the optimal feedback, if any, is identical for both problems. But, because the update of $f(.)$ in (39) is linear, the derivative of the term inside the inf of (39) with respect to y is easier to obtain. Hence, finding the optimal y is easier with (39) than it is with (29). For this reason, we shall work with (39) and (40).

6.2. Steps of the Proof of Theorem 1

The proof of Theorem 1 is based on the monotonicity argument, which is a classical tool for variational inequalities, quasi-variational inequalities and Bellman equations at large. We adopt this argument to our case of partially observed demand to complete the proof of Theorem 1 given in EC.4. The proof requires the following four steps:

• Since $W(x,f(.))$ must be positive, we will see that it is enough to require that instead of the sharper constraint on the left of (40), and consider the set Γ of functionals on $\mathbb{R}\times L^1_{\varphi+}(\mathbb{R})$ such that

$$\Gamma=\Big\{W(x,f(.))|\ W \text{ is continuous}, \quad 0\le W(x,f(.))\le\big(\frac{hx^+}{1-\alpha}+(p+c)x^-\big)\int f(\eta)d\eta+\frac{\alpha(p+c)}{1-\alpha}\int\varphi(\eta)f(\eta)d\eta\Big\}, \tag{41}$$

and define a monotone nonlinear operator T on Γ as follows:

$$T(W)(x,f(.))=(hx^++px^--cx)\int f(\eta)d\eta+\inf_{y\in\mathcal{L}_{x,f(.)}}\big[cy\int f(\eta)d\eta+\alpha\int_0^{+\infty}W(y-z,f(.)g(z|.))dz\big]. \tag{42}$$

• We show that the operator T maps Γ into itself, so $T(W)(x,f(.))$ is continuous in both arguments x and $f(.)$. For the proof, we first show that $T(W)(x,f(.))\in\Gamma$. As for the continuity, observe that $W(x,f(.))$ is continuous by (41), so the function of y in the second term inside the infimum in (42) is continuous, and thus it is minimized at y on the bounded interval $\mathcal{L}_{x,f(.)}$. Then, by the measurable selection theorem, we can find a Borel function $\hat{y}(x,f(.))$ that realizes the infimum in (42) for any $(x,f(.))\in\mathbb{R}\times L^1_{\varphi+}(\mathbb{R})$, i.e.,

$$\left|\begin{array}{l} T(W)(x,f(.))=(hx^++px^--cx)\int f(\eta)d\eta\ + \\ c\hat{y}(x,f(.))\int f(\eta)d\eta+\alpha\int_0^{+\infty}W(\hat{y}(x,f(.))-z,f(.)g(z|.))dz,\ \ \forall x,f(.), \\ \\ \hat{y}(x,f(.))\in\mathcal{L}_{x,f(.)}. \end{array}\right.$$

Next, we consider $x_n,\ x \in \mathbb{R},\ x_n \to x$ and $f_n(.),\ f(.) \in L^1_{\varphi+}(\mathbb{R}),\ f_n(.) \to f(.) \in L^1_{\varphi+}(\mathbb{R})$. We then prove that the sequence $T(W)(x_n, f_n(.))$ converges to $T(W)(x, f(.))$, which proves the continuity of $T(W)(x, f(.))$.

• For the existence of a solution, we first note that (39) is a fixed point equation. Then, the functional $W(x, f(.))$ must satisfy $W(x, f(.)) = T(W)(x, f(.))$. We define two monotone sequences:

$$\left| \begin{array}{l} W_{k+1}(x, f(.)) = T(W_k)(x, f(.)), \\ \\ W_0(x, f(.)) = 0, \end{array} \right.$$

$$\left| \begin{array}{l} W^{k+1}(x, f(.)) = T(W^k)(x, f(.)), \\ W^0(x, f(.)) = (\dfrac{hx^+}{1-\alpha} + (p+c)x^-) \displaystyle\int f(\eta)d\eta + \dfrac{\alpha(p+c)}{1-\alpha} \displaystyle\int \varphi(\eta) f(\eta) d\eta, \end{array} \right.$$

and show that $W_k(x, f(.))$ is lower semi-continuous and it converges to a lower semi-continuous function $\underline{W}(x, f(.))$ and $W^k(x, f(.))$ is upper semi-continuous and it converges to an upper semi-continuous function $\overline{W}(x, f(.))$. In fact, $\underline{W}(x, f(.))$ is the smallest solution and $\overline{W}(x, f(.))$ is the largest solution, in the sense that if $W(x, f(.))$ is any solution of (39), then necessarily

$$\underline{W}(x, f(.)) \leq W(x, f(.)) \leq \overline{W}(x, f(.)). \tag{43}$$

• For the uniqueness of the solution, we go back to functions $f(.)$ that are probability densities. We can transpose the results in W to results in Φ, and thus we obtain the minimum and maximum solutions $\underline{\Phi}(x, f(.))$ and $\overline{\Phi}(x, f(.))$, respectively, of (29). Thanks to the feedback $\hat{y}(x, f(.))$ given by the selection theorem, we construct the processes $\hat{x}_n$, $\hat{y}_n$, and $\hat{f}_n(.)$ by the iterations

$$\hat{y}_n = \hat{y}(\hat{x}_n, \hat{f}_n(.)), \tag{44}$$

$$\hat{x}_{n+1} = \hat{y}_n - D_n, \quad \hat{x}_1 = x, \tag{45}$$

$$\hat{f}_{n+1}(\xi) = \hat{f}_n(\xi) \frac{g(D_n|\xi)}{\int g(D_n|\eta)\hat{f}_n(\eta)d\eta}, \quad \hat{f}_1(.) = f(.). \tag{46}$$

We show that the minimum solution coincides with the value function by first setting $\hat{V} = (\hat{y}_1, \cdots, \hat{y}_n, \cdots)$ and showing that

$$\underline{\Phi}(x, f(.)) \geq J_{x,f(.)}(\hat{V}) \geq \Phi^*(x, f(.)) = \inf_V J_{x,f(.)}(V).$$

Since $\underline{\Phi}(x, f(.))$ is the smallest solution, it follows that all solutions are larger than the value function. On the other hand, by Proposition 1, all solutions are smaller than the value function. Necessarily, the solution is unique and coincides with the value function. Moreover, the feedback $\hat{y}(x, f(.))$ allows us to construct an optimal control by using (44)-(46). That this feedback is a base stock policy will be proved as the main result of the next section.

Before proceeding to the next section, let us mention that the following result follows from the facts that $\underline{\Phi}(x, f(.))$ is l.s.c., $\overline{\Phi}(x, f(.))$ is u.s.c., and $\underline{\Phi}(x, f(.)) = \overline{\Phi}(x, f(.)) = \Phi(x, f(.))$.

Corollary 1. *The solution $\Phi(x, f(.))$ obtained in Theorem 1 is continuous.*

7. Optimality of Base Stock Policy

We have seen in Section 3 that when there is no unknown parameter, there is no learning, and the optimal feedback is a base stock policy defined in (7). The objective here is to obtain a similar result when there is learning, with a base stock now depending on the latest belief function updated from observing the demand realizations. It is convenient to consider an equivalent problem by setting

$$Z(x, f(.)) = W(x, f(.)) - (hx^+ + px^- - cx) \int f(\eta) d\eta. \tag{47}$$

Let us note that we do not need to impose the upper bound on y for minimizing the right hand side of (47), as it will be automatically satisfied at a minimum point.

Lemma 2. *The Bellman equation (47) can be written as*

$$\begin{aligned} Z(x, f(.)) =& \alpha(c-h) \int \varphi(\eta) f(\eta) d\eta + \min_{y \geq x} \Big[(c(1-\alpha) + \alpha h) y \int f(\eta) d\eta \\ & + \alpha(h+p) \int_y^{+\infty} \int \bar{G}(z|\eta) f(\eta) d\eta dz + \alpha \int_0^{+\infty} Z(y-z, f(.)g(z|.)) dz \Big]. \end{aligned} \tag{48}$$

This transformation is justified in EC.5. The advantage of this formulation is that x appears only in the constraint on y. We know from Theorem 1 and Corollary 1 that this equation has a unique continuous solution in the following interval:

$$
\begin{aligned}
Z_0(x,f(.)) &\stackrel{\text{def}}{=} (\frac{\alpha h x^+}{1-\alpha}+cx)\int f(\eta)d\eta - \frac{\alpha h}{(1-\alpha)^2}\int \varphi(\eta)f(\eta)d\eta \\
&\leq Z(x,f.)) \\
&\leq Z^0(x,f(.)) \stackrel{\text{def}}{=} (\frac{\alpha h}{1-\alpha}+c)x^+\int f(\eta)d\eta + \frac{\alpha(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta.
\end{aligned}
\tag{49}
$$

The solution of (48) is a fixed point of the operator

$$
Z(x,f(.)) = \Theta(Z)(x,f(.)) \tag{50}
$$

with

$$
\begin{aligned}
\Theta(Z)(x,f(.)) \stackrel{\text{def}}{=} & \alpha(c-h)\int \varphi(\eta)f(\eta)d\eta + \min_{y\geq x}\Big[(c(1-\alpha)+\alpha h)y\int f(\eta)d\eta \\
&+\alpha(h+p)\int_y^{+\infty}\int \bar{G}(z|\eta)f(\eta)d\eta dz + \alpha\int_0^{+\infty} Z(y-z,f(.)g(z|.))dz\Big],
\end{aligned}
\tag{51}
$$

which is a nonlinear operator on functionals $Z(x,f(.))$ defined on $\mathbb{R}\times L^1_{\varphi+}(\mathbb{R})$ and satisfying the bounds in (49).

7.1. Preserving Convexity

We state the following property whose proof is given in EC.6. This result helps us to establish the stability needed in the proof of the optimality of a base stock policy.

PROPOSITION 2. *Assume that $x\to Z(x,f(.))$ is differentiable, increasing and convex in x. Assume that*

$$
\sup_{x,f(.)}\frac{|Z'(x,f(.))|}{\int f(\eta)d\eta} < +\infty \tag{52}
$$

and

$$
-(\alpha p - c(1-\alpha))\int f(\eta)d\eta + \alpha\int_0^{+\infty} Z'(y-z,f(.)g(z|.))dz < 0, \tag{53}
$$

where $Z'(x,f(.))$ is the derivative of $Z(x,f(.))$ with respect to x. Then $x\to\Theta(Z)(x,f(.))$ and $x\to Z(x,f(.))$ have the same properties.

7.2. Base Stock Policy

It is interesting to find $Z'(x,f(.))$ directly, and not by differentiating $Z(x,f(.))$. Indeed, we have

$$Z'(x,f(.)) = \left| \begin{array}{l} 0, \quad \text{if } \; x < S_Z(f(.)), \\ \\ (c(1-\alpha)+\alpha h)\displaystyle\int f(\eta)d\eta - \alpha(h+p)\int \bar{G}(x|\eta)f(\eta)d\eta \\ +\alpha \displaystyle\int_0^{+\infty} Z'(x-z,f(.)g(z|.))dz, \quad \text{if } \; x \geq S_Z(f(.)). \end{array} \right. \tag{54}$$

Since the second expression is positive, $Z'(x,f(.))$ satisfies

$$\begin{aligned} Z'(x,f(.)) = \Big((c(1-\alpha)+\alpha h)\int f(\eta)d\eta - \alpha(h+p)\int \bar{G}(x|\eta)f(\eta)d\eta \\ +\alpha \int_0^{+\infty} Z'(x-z,f(.)g(z|.))dz \Big)^+, \end{aligned} \tag{55}$$

and we define

$$S_Z(f(.)) = \sup\{x \;\mid\; Z'(x,f(.)) = 0\}. \tag{56}$$

We can now state the base stock policy in the following theorem whose proof is relegated to EC.7.

THEOREM 2. *We make the assumptions of Theorem 1. The function $Z(x,f(.))$ is globally continuous and $\mathcal{C}^1$, increasing, and convex in x. The optimal feedback $\hat{y}(x,f(.))$ takes the following base stock policy form:*

$$\hat{y}(x,f(.)) = \left| \begin{array}{l} S_Z(f(.)), \quad \text{if} \quad x < S_Z(f(.)), \\ \\ x, \quad \text{if} \quad x \geq S_Z(f(.)). \end{array} \right. \tag{57}$$

We have the following result whose proof is given in EC.7.

PROPOSITION 3. *Equation (55) has one and only one solution in the functional space of continuous functions on $\mathbb{R} \times L^1_{\varphi+}(\mathbb{R})$ with the norm*

$$\sup_{x,f(.)} \frac{|Z'(x,f(.))|}{\int f(\eta)d\eta} < +\infty. \tag{58}$$

Knowing $Z'(x,f(.))$ and $S_Z(f(.))$, we recover $Z(x,f(.))$ by

$$Z(x,f(.)) = \left| \begin{array}{l} K(f(.)), \quad \text{if } x < S_Z(f(.)), \\ \\ K(f(.)) + \displaystyle\int_{S_Z(f(.))}^{x} Z'(\eta,f(.))d\eta, \quad \text{if } x \geq S_Z(f(.)). \end{array} \right. \tag{59}$$

It remains to obtain $K(f(.))$, the solution of a linear problem derived from (48). Namely,

$$\begin{aligned} K(f(.)) &= \alpha(c-h)\int \varphi(\eta)f(\eta)d\eta + (c(1-\alpha)+\alpha h)S_Z(f(.))\int f(\eta)d\eta \\ &+ \alpha(h+p)\int_{S_Z(f(.))}^{+\infty}\int \bar{G}(z|\eta)f(\eta)d\eta\, dz \\ &+ \alpha\int_0^{+\infty}\Big(\int_{-\infty}^{S_Z(f(.))-z} Z'(\eta,f(.)g(z|.))d\eta\Big)dz + \alpha\int_0^{+\infty} K(f(.)g(z|.))dz. \end{aligned} \tag{60}$$

Note that since $K(f(.)) = Z(0,f(.)) = W(0,f(.))$, we can use the right inequality of (49) to have

$$0 \leq K(f(.)) \leq \frac{\alpha(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta. \tag{61}$$

7.3. A Monotonicity Property

We consider the iterative sequence

$$\begin{aligned} Z'_{k+1}(x,f(.)) = \Big(&(c(1-\alpha)+\alpha h)\int f(\eta)d\eta - \alpha(h+p)\int \bar{G}(x|\eta)f(\eta)d\eta \\ &+ \alpha\int_0^{+\infty} Z'_k(x-z,f(.)g(z|.))dz\Big)^+, \text{ with } Z'_0(x,f(.)) = 0. \end{aligned} \tag{62}$$

By recurrence, we see that

$$Z'_k(x,f(.)) \geq 0, \text{ and increasing in } x. \tag{63}$$

Consequently,

$$(c(1-\alpha)+\alpha h)\int f(\eta)d\eta - \alpha(h+p)\int \bar{G}(x|\eta)f(\eta)d\eta + \alpha\int_0^{+\infty} Z'_k(x-z,f(.)g(z|.))dz$$

is increasing in x. Since it is $-(\alpha p - c(1-\alpha))\int f(\eta)d\eta$ at $x=-\infty$ and larger than $(c(1-\alpha)+$

$\alpha h)\displaystyle\int f(\eta)d\eta$ at $x=+\infty$, there exists a unique $S_{Z_{k+1}}(f(.))$ such that

$$Z'_{k+1}(x,f(.))=0,\ \text{ if }\ x\leq S_{Z_{k+1}}(f(.)). \tag{64}$$

Also $Z'_1(x,f(.))\geq Z'_0(x,f(.))$. By recurrence again, we have

$$Z'_{k+1}(x,f(.))\geq Z'_k(x,f(.)). \tag{65}$$

Since (55) is a fixed point for a contraction, we can assert that

$$Z'_k(x,f(.))\uparrow Z'(x,f(.)). \tag{66}$$

From the definition of $S_{Z_k}(f(.))$, we get immediately

$$S_{Z_{k+1}}(f(.))\leq S_{Z_k}(f(.)) \tag{67}$$

and

$$S_{Z_k}(f(.))\uparrow S_Z(f(.)). \tag{68}$$

8. Learning with Conjugate Probabilities

In this section we take the important case of conjugate probabilities. Most of the literature devoted to proving the optimality of the base stock policy in the presence of demand learning, including Iglehart (1964) and Azoury (1985), follow Scarf (1959) and Scarf (1960a) and assume a conjugate probability set up. In these papers, the analysis is done via a sequence of functions. By contrast, we do not need that set up as we recover the results as a particular case of our general theory. Here we specialize the belief function $f(.)$ introduced in our general theory to depend on a vector X of hyperparameters. So we write it as $f_X(\xi)$. We shall then see that the infinite dimensional problem of the previous sections reduces to a finite dimensional one, the dimension being the size of the vector of the hyperparameters. Our analysis therefore generalizes as well as rigorizes the literature mentioned above.

In Subsection 8.1, we consider a special case of Weibull demand whose scale parameter is not known. So we choose our belief function from its conjugate prior family of Gamma densities, which is characterized by two hyperparameters. Subsection 8.2 shows that the general problem gives the hyperparameter vector X as the sufficient statistics and provides an equation to be satisfied by the base stock expressed as a function of X. In Subsection 8.3, we apply this analysis to the Weibull demand case of Subsection 8.1, and obtain the base stock as a function of the two Gamma hyperparameters. We also apply the developed theory to two numerical examples to illustrate its implementation. That is, we simulate demands according to supposedly true exponential and Weibull densities, obtain the sequence of optimal base stocks as learned from the simulated demands, order accordingly, and evolve the inventory dynamics in each of the two examples. We also validate the model by simulating demands for one thousand periods and see that the sequence base stocks is tending to converge to the true base stocks in each case.

We set

$$b(z,X) = \frac{\int f_X(\eta) g(z|\eta) d\eta}{\int f_X(\eta) d\eta}, \tag{69}$$

which defines a probability density. We assume that

$$\frac{f_X(\xi) g(z|\xi)}{\int f_X(\eta) d\eta} = b(z,X) \frac{f_{H(z,X)}(\xi)}{\int f_{H(z,X)}(\eta) d\eta}, \tag{70}$$

which defines a coupling between the family $f_X(\xi)$ and the probability density $g(z|\xi)$. $H(z,X)$ is a vector of same size as X. These are known as conjugate probabilities.

8.1. Example of Weibull Demand

As a special example, we consider the demand to be distributed according to Weibull probability density

$$g(z|\xi) = m\xi z^{m-1} e^{-\xi z^m}, \tag{71}$$

which was also studied in Azoury (1985) and Bensoussan (2011). In this parametrization, m is the shape parameter and $\xi^{-1/m}$ is the scale parameter. The mean demand is $\xi^{-1/m}\Gamma(1+1/m)$, and thus as ξ decreases, the mean demand increases. As is common in the related literature, we assume that the shape parameter m is known and ξ is not known. Note that when $m=1$, (71) reduces to the exponential density with the mean given by $1/\xi$.

As for our belief of ξ, we will consider the Gamma family $f_X(\xi)$ given by

$$f_X(\xi) = \xi^{a-1}e^{-\lambda\xi}, \tag{72}$$

with the hyperparameter $X=(\lambda,a)$, $a>0$, where λ and a are known as the rate and shape hyperparameters, respectively, in this shape-rate parametrization of the unnormalized Gamma density. For this unnormalized version of the Gamma density, $\int f_X(\eta)d\eta = \frac{\Gamma(a)}{\lambda^a}$, where $\Gamma(a) = \int_0^{+\infty} x^{a-1}e^{-x}\,dx$ is the Gamma function and the normalized version is $\frac{\lambda^a f_X(\xi)}{\Gamma(a)}$. Furthermore, a reflects the number of prior observations, λ denotes the sum of prior observations, and the mean is given by a/λ.

One can also see from (15) and (72) that the posterior Gamma density

$$f_{n+1}(\xi) = \frac{\left(\lambda + \sum_{i=1}^n D_i^m\right)^{a+n}\xi^{a+n-1}e^{-(\lambda+\sum_{i=1}^n D_i^m)\xi}}{\Gamma(a+n)} \tag{73}$$

with its mean given by $\frac{a+n}{\lambda+\sum_{i=1}^n D_i^m}$. Substituting (69) in (70), we get

$$\frac{f_X(\xi)g(z|\xi)}{\int f_X(\eta)g(z|\eta)d\eta} = \frac{f_{H(z,X)}(\xi)}{\int f_{H(z,X)}(\eta)d\eta}, \tag{74}$$

and substituting (71) and (72) in (69) yields the conditional demand density

$$b(z,X) = \frac{ma\lambda^a z^{m-1}}{(\lambda+z^m)^{a+1}}, \tag{75}$$

given the hyperparameter $X=(\lambda,a)$, with the mean demand given by

$$\frac{\lambda^{1/m}a\Gamma(a-1/m)\Gamma(1+1/m)}{\Gamma(a+1)}, \quad ma>1. \tag{76}$$

Note that a change of variable $x = \frac{z}{\lambda^{1/m}}$ applied to (75) gives $b(x,X) = \frac{max^{m-1}}{(1+x^m)^{a+1}}$, which is Burr Type XII distribution, with CDF $B(x,X) = 1 - \frac{1}{(1+x^m)^a}$. This distribution will be useful in reducing the dimension of the hyperparameter from (λ, a) to just a, as we will see later when we reduce the integral equation (86) to (88).

The updated demand is (first-order) stochastically decreasing in a and stochastically increasing in λ. Thus, both hyperparameters are measures of the predicted size of the underlying demand. However, since the shape hyperparameter reflects the number of observations, it is the sole measure of the precision with which the underlying demand is known. Indeed, as the number of demand observations increases, the uncertainty about the unknown parameter ξ in the demand density decreases. In the limit, while demand remains stochastic, ξ is known exactly. Using (71)-(75) in (70) we can see that it is satisfied with $H(z,X) = (\lambda + z^m, a+1)$, i.e., $f_{H(z,X)}(\xi) = \xi^a e^{-(\lambda+z^m)\xi}$.

8.2. Sufficient Statistics

Back to the general case, we want to study equation (55) for conjugate probabilities. We shall see that our infinite dimensional problem reduces to a problem of the dimension equals to the size of the vector X of hyperparameters. We first note that $Z'(x, f(.))$ has the same property as $W(x, f(.))$ in (37), namely,

$$Z'(x, \mu f(.)) = \mu Z'(x, f(.)), \ \ \forall \mu > 0. \tag{77}$$

We write (55) as

$$\begin{aligned} Z'(x, f_X(.)) = \Big((c(1-\alpha) + \alpha h) \int f_X(\eta) d\eta - \alpha(h+p) \int \bar{G}(x|\eta) f_X(\eta) d\eta \\ + \alpha \int_0^{+\infty} Z'(x-z, f_X(.)g(z|.))dz \Big)^+. \end{aligned} \tag{78}$$

Using the property (77), we can write

$$\begin{aligned} Z'(x, \frac{f_X(.)}{\int f_X(\eta)d\eta}) = \Bigg((c(1-\alpha) + \alpha h) - \alpha(h+p) \frac{\int \bar{G}(x|\eta) f_X(\eta) d\eta}{\int f_X(\eta) d\eta} \\ + \alpha \int_0^{+\infty} Z'\Big(x - z, \frac{f_X(.)g(z|.)}{\int f_X(\eta) g(z|\eta) d\eta}\Big) \frac{\int f_X(\eta) g(z|\eta) d\eta}{\int f_X(\eta) d\eta} dz \Bigg)^+. \end{aligned} \tag{79}$$

Then we let

$$R(x,X)=Z'\Big(x,\frac{f_X(.)}{\int f_X(\eta)d\eta}\Big) \tag{80}$$

and

$$\bar{B}(x,X)=\frac{\int \bar{G}(x|\eta)f_X(\eta)d\eta}{\int f_X(\eta)d\eta}=\int_x^{+\infty}b(z,X)dz. \tag{81}$$

The important point is to note that

$$\frac{f_X(\xi)g(z|\xi)}{\int f_X(\eta)g(z|\eta)d\eta}=\frac{f_{H(z,X)}(\xi)}{\int f_{H(z,X)}(\eta)d\eta}. \tag{82}$$

Therefore, (79) becomes

$$R(x,X)=\Big((c(1-\alpha)+\alpha h-\alpha(h+p)\bar{B}(x,X)+\alpha\int_0^{+\infty}R(x-z,H(z,X))b(z,X)dz\Big)^+. \tag{83}$$

This is a finite dimensional functional equation. It has one and only one solution on the set of bounded functions of x and X. We have $R(x,X)=0$ for $x\leq 0$, and there exists a single $S_R(X)>0$ such that $R(x,X)=0$ for $x\leq S_R(X)$. Moreover, $S_R(X)$ can be obtained as the unique solution of

$$\Big(c(1-\alpha)+\alpha h)-\alpha(h+p)\bar{B}(S_R(X),X)+\alpha\int_0^{+\infty}R(S_R(X)-z,H(z,X))b(z,X)dz\Big)^+=0. \tag{84}$$

8.3. Computation of Base Stock for Weibull Demand

We consider the Weibull demand example given by (71) and (72). We have $X=(\lambda,a)$ and $H(z,X)=(\lambda+z^m,a+1)$, and therefore using (75) in (81) and (84), we have

$$\bar{B}(z,X)=\Big(\frac{\lambda}{\lambda+z^m}\Big)^a, \tag{85}$$

and

$$R(x,\lambda,a)=\Big(c(1-\alpha)+\alpha h-\alpha(h+p)\Big(\frac{\lambda}{\lambda+x^m}\Big)^a+\alpha\int_0^x R(x-z,\lambda+z^m,\,a+1)\frac{ma\lambda^a z^{m-1}}{(\lambda+z^m)^{a+1}}dz\Big)^+, \tag{86}$$

$x>0$. We find easily that

$$R(x,\lambda,a)=U_a\Big(\frac{x}{\lambda^{\frac{1}{m}}}\Big), \tag{87}$$

with $U_a(x)$ satisfying

$$U_a(x)=\Big(c(1-\alpha)+\alpha h-\alpha(h+p)\Big(\frac{1}{1+x^m}\Big)^a+\alpha ma\int_0^x U_{a+1}\Big(\frac{x-\zeta}{(1+\zeta^m)^{\frac{1}{m}}}\Big)\frac{\zeta^{m-1}}{(1+\zeta^m)^{a+1}}d\zeta\Big)^+. \tag{88}$$

We have $U_a(x) = 0$ for $x \leq 0$. We have thus reduced the dimensionality of the problem from (x, λ, a) to just (x, a) and have obtained a simpler recursion (88), as is common in the related literature. From (84) we see that the base stock $S_U(\lambda, a) = \lambda^{1/m} L_U(a)$, where $L_U(a)$ is the solution of

$$c(1-\alpha) + \alpha h - \alpha(h+p)\Big(\frac{1}{1+L_U^m(a)}\Big)^a + \alpha m a \int_0^{L_U(a)} U_{a+1}\Big(\frac{L_U(a)-\zeta}{(1+\zeta^m)^{\frac{1}{m}}}\Big)\frac{\zeta^{m-1}}{(1+\zeta^m)^{a+1}} d\zeta = 0. \quad (89)$$

This allows managers to use (88) and (89) to first compute in advance a surrogate base stock $L_U(a)$ for each period that depends only on the shape hyperparameter a of that period and not on the scale (size) of the demand. Then determine the optimal base stock $\lambda^{1/m} L_U(a)$ to use in each period by scaling the surrogate by the current best estimate of the demand, given by the scale hyperparameter consisting of the sum of the initial belief of the demand and the observed past demands, each raised to the power m.

We approximate the solution of (88) by the recursive scheme

$$U_a^k(x) = \Big(c(1-\alpha) + \alpha h - \alpha(h+p)\Big(\frac{1}{1+x^m}\Big)^a + \alpha m a \int_0^x U_{a+1}^{k-1}\Big(\frac{x-\zeta}{(1+\zeta^m)^{\frac{1}{m}}}\Big)\frac{\zeta^{m-1}}{(1+\zeta^m)^{a+1}} d\zeta\Big)^+, \quad (90)$$

for $k = 1, 2, \cdots$, with the initial guess $U_{a+1}^0(x) = 0$. The solution $U_a(x)$ is given by $\lim_{k\to\infty} U_a^k(x)$. We illustrate the procedure numerically for the parameter values $\alpha = 0.9$, $p = 4$, $c = 10$, $h = 1$, and $m = 1$ and 2.

8.3.1. Case $m = 1$, $g(z|\xi) = \xi e^{-\xi z}$

We carry out our computations for $a = 1, \cdots 6$, but display the results for only $a = 1$ and $a = 2$ in Figure 1 and Figure 2, respectively.

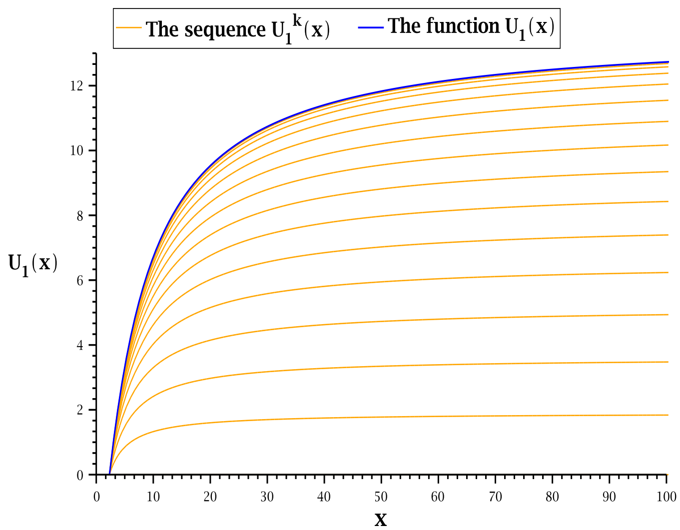

Figure 1 **Convergence to the function** $U_1(x)$.

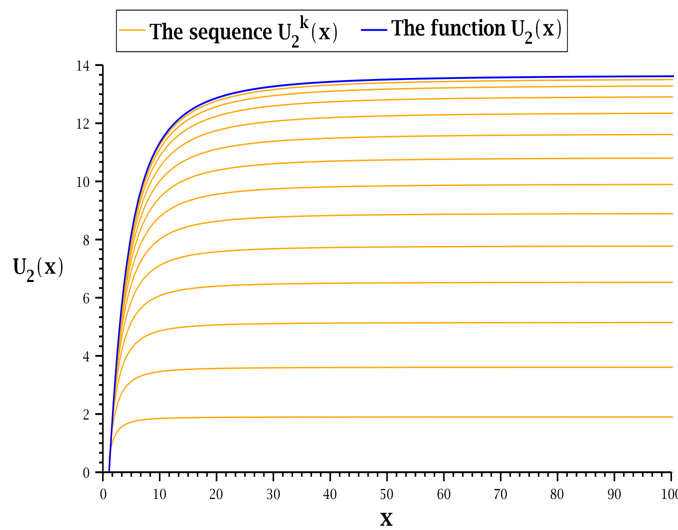

Figure 2 **Convergence to the function** $U_2(x)$

Once $U_1(x)$, $U_2(x), \cdots, U_6(x)$ are obtained, (89) can be used to find $L_U(1) = 2.316$, $L_U(2) = 0.823$, $L_U(3) = 0.493$, $L_U(4) = 0.351$, and $L_U(5) = 0.272$. So, the formula $S_U(\lambda, a) = \lambda^{1/m} L_U(a)$ yields the base stocks $S_U(\lambda_1, 1) = 2.316\lambda_1$, $S_U(\lambda_2, 2) = 0.823\lambda_2$, $S_U(\lambda_3, 3) = 0.493\lambda_3$, $S_U(\lambda_4, 4) = 0.351\lambda_4$, and $S_U(\lambda_5, 5) = 0.272\lambda_5$.

To illustrate how to implement our model, let us suppose that the true demand, which is exponential when $m = 1$, has $\xi = 0.25$, and thus has 4 as its mean. Then, the true base stock as computed using (6) is $S_1 = 4.795$. However, the IM does not know this value, and so he must make an initial guess of the hyperparamater X. We will illustrate with two different guesses: $X_1 = (\lambda_1, a_1) = (3, 1)$ and $X_2 = (\lambda_1, a_1) = (5, 1)$, respectively, where λ_1/a_1 represents the mean value of ξ according to the conjugate Gamma distribution.

Next, we use the predefined Matlab function *exprnd*$(1/\xi, 1, N)$ to generate $N = 5$ simulated demands from the exponential distribution with $\xi = 0.25$. These are $D_1 = 6.418$, $D_2 = 0.409$, $D_3 = 0.458$, $D_4 = 3.895$, and $D_5 = 5.542$.

We first start with our guess of the hyperparameter $X_1 = (3,1)$ and consider the inventory evolution (14) with, say, the initial inventory $x_1 = 0$. We compute the base stock $S_U(3,1) = 2.316\lambda_1 = 6.948$, and so we order up to it to obtain the order quantity $v_1 = 6.948$, which gives $x_2 = x_1 + v_1 - D_1 = 0.530$. In period 2, $(\lambda_2, a_2) = (\lambda_1 + D_1, a_1 + 1) = (9.418, 2)$, and the base stock $S_U(\lambda_2, a_2) = 0.823\lambda_2 = 7.751$. Thus, the order quantity $v_2 = 7.221$ and $x_3 = x_2 + v_2 - D_2 = 7.342$. Following this process, we obtain the inventory dynamics for the first five periods presented in Table 1. Similarly, we obtain Table 2 with a different initial guess $X_2 = (\lambda_1, a_1) = (5,1)$.

Table 1 **Optimal path starting with** $X_1 = (3,1)$.

D_n	6.418	0.409	0.458	3.895	5.542
a_n	1	2	3	4	5
λ_n	3	9.418	9.827	10.285	14.180
$S_U(\lambda_n, a_n)$	6.948	7.751	4.845	3.610	3.857
v_n	6.948	7.221	0	0	0.868
x_n	0	0.530	7.342	6.884	2.989

Table 2 **Optimal path starting with** $X_2 = (5,1)$.

D_n	6.418	0.409	0.458	3.895	5.542
a_n	1	2	3	4	5
λ_n	5	11.418	11.827	12.285	16.180
$S_U(\lambda_n, a_n)$	11.580	9.397	5.831	4.312	4.401
v_n	11.580	4.235	0	0	0
x_n	0	5.162	8.988	8.530	4.635

In the fourth row in each table, we see that the base stock appears to be moving toward its true value of 4.795.

We also run the simulation for $N = 1000$ periods and compute $L_U(a_N) = 0.0012$. For each of the two initial guesses $\lambda_1 = 3$ and 5, we obtain $\lambda_N = \lambda_1 + \sum_{i=1}^{N} D_i = 3988.4$ and 3990.4 and the respective base stocks $S_U(\lambda_N, a_N) = 0.0012\lambda_N = 4.786$ and 4.788. As we can, these values are very close to the true base stock of 4.795.

8.3.2. Case $m = 2$, $g(z|\xi) = 2\xi z e^{-\xi z^2}$

We now suppose the true demand to be Weibull with $m = 2$ and $\xi = 0.25$. The mean demand is 1.772, and the true base stock as computed using (6) is $S_2 = 2.190$. As before, the IM makes two initial guesses of the hyperparameter: $X_1 = (\lambda_1, a_1) = (3,1)$ and $X_2 = (\lambda_1, a_1) = (5,1)$, respectively, where we note that λ_1/a_1 represents the mean value of ξ according to the conjugate Gamma distribution. We use the same technique to obtain $S_U(\lambda_1,1)=1.522\lambda_1$, $S_U(\lambda_2,2)=0.910\lambda_2$, $S_U(\lambda_3,3)=0.701\lambda_3$, $S_U(\lambda_4,4)=0.593\lambda_4$, and $S_U(\lambda_5,5)=0.521\lambda_5$. Now the method for generating Weibull random demands is based on the standard inverse transform technique. These are $D_1=2.769$, $D_2=1.315$, $D_3=1.973$, $D_4=1.697$, and $D_5=1.175$. We present the inventory dynamics for the first five periods in Table 3 and Table 4 corresponding to the two initial guesses.

Table 3 **Optimal path starting with** $X_1=(3,1)$.

D_n	2.769	1.315	1.973	1.697	1.175
a_n	1	2	3	4	5
λ_n	3	10.667	12.396	16.289	19.169
$S_U(\lambda_n,a_n)$	2.636	2.972	2.468	2.393	2.281
v_n	2.636	3.105	0.811	1.898	1.585
x_n	0	-0.133	1.657	0.495	0.696

Table 4 **Optimal path starting with** $X_2=(5,1)$.

D_n	2.769	1.315	1.973	1.697	1.175
a_n	1	2	3	4	5
λ_n	5	12.667	14.396	18.289	21.169
$S_U(\lambda_n,a_n)$	3.403	3.239	2.660	2.536	2.397
v_n	3.403	2.604	0.736	1.849	1.558
x_n	0	0.634	1.923	0.686	0.838

From the fourth row in each table, we see that the base stock appears to be moving toward its true value of 2.190.

As in Subsection 8.3.1, we also run the simulation for $N=1000$ periods and compute $L_U(a_N)=0.034$. For each of the two initial guesses $\lambda_1=3$ and 5, we obtain $\lambda_N=\lambda_1+\sum_{i=1}^{N}D_i^2=4164.6$ and 4166.6 and the respective base stocks $S_U(\lambda_N,a_N)=0.034\lambda_N^{1/2}=2.194$ and 2.195. As we can see, these values are very close to the true base stock of 2.190.

9. When Demand Follows One of Two Distributions

In this section we consider the special case when the demand is known to come from one of two possible distributions, but we don't know which. An example would be high demand and low demand environments. Thus, we model this in Subsection 9.1 by a belief function with a two-dimensional hyperparameter, which can be reduced in this case to simply their ratio, and rewrite the functional equation for a special case of exponential demand. This case affords us a possibility of convenient approximations. In Subsections 9.1.1 and 9.1.3, we develop two approximation procedures involving the iteration of the unique solution of the functional equation and the iteration by piecewise constant functions, respectively. We show in EC.10 that the two approximations produce results that are very close. In Subsection 9.1.2 we illustrate the method in Subsection 9.1.1 by applying it to an example of high and low exponential demands.

We should mention that the methods developed here can handle very general demand distributions including nonparametric ones. For example, we can have bimodal distributions that cannot be treated in the conjugate probability framework of Section 8. Moreover, the methodology can be easily generalized to permit a finite number of possible general distributions.

Let us consider the belief function

$$f(\xi) = \beta_0\delta(\xi) + \beta_1\delta(\xi - 1), \tag{91}$$

where $\delta(\xi)$ and $\delta(\xi-1)$ represent the Dirac masses at 0 and 1, respectively, and β_0 and β_1 are hyperparameters. Since it is not a probability density, so this writing is formal. We set the demands $g_0(z) = g(z|0)$ and $g_1(z) = g(z|1)$, which can be very general. We have

$$g(z|\xi)f(\xi) = g_0(z)\beta_0\delta(\xi) + g_1(z)\beta_1\delta(\xi - 1), \tag{92}$$

which gives us a formal representation the demand measure given the values of the hyperparameters β_0 and β_1. Note that it preserves, in relation to $f(\xi)$, the fact that it is also the sum of two Dirac measures at 0 and 1, respectively. The functional $Z'(x, \beta_0\delta(\xi) + \beta_1\delta(\xi - 1))$ in (55) can be written as $Z'(x, \beta_0, \beta_1)$, depending on the hyperparameters β_0 and β_1 instead of $f(.)$, and it satisfies the equation

$$\begin{aligned} Z'(x, \beta_0, \beta_1) = \Big(&(c(1-\alpha) + \alpha h)(\beta_0 + \beta_1) - \alpha(h+p)(\beta_0\bar{G}_0(x) + \beta_1\bar{G}_1(x)) \\ &+ \alpha \int_0^{+\infty} Z'\Big(x - z, \beta_0 g_0(z), \beta_1 g_1(z)\Big)dz\Big)^+. \end{aligned} \tag{93}$$

We can assume both β_0 and β_1 to be strictly positive and set $\beta = \beta_0/\beta_1$, since if we begin with one of them to be zero then we know the true demand and there is no need of learning. Since $\beta_0 + \beta_1 = 1$, we have

$$\beta_0 = \beta/(\beta+1) \text{ and } \beta_1 = 1/(\beta+1). \tag{94}$$

Then noting (77), we have $Z'(x,\beta_0,\beta_1) = Z'(x,\beta\beta_1,\beta_1) = Z'(x,\beta,1)\beta_1$. Then by setting $Z'(x,\beta) = Z'(x,\beta,1)$, we obtain immediately from (93) the equation

$$\begin{aligned} Z'(x,\beta) = \Big(&(c(1-\alpha)+\alpha h)(\beta+1) - \alpha(h+p)(\beta\bar{G}_0(x)+\bar{G}_1(x)) \\ &+ \alpha\int_0^{+\infty} g_1(z) Z'\Big(x-z, \beta\frac{g_0(z)}{g_1(z)}\Big)dz\Big)^+. \end{aligned} \tag{95}$$

Thus, we are left only with one hyperparameter β to learn from the demand observations.

The remainder of Section 9 will be devoted to solving this equation by using approximation methods. For convenience in exposition, we will only consider the case of exponential demands, and develop two approximations in Subsections 9.1.1 and 9.1.3, respectively. In Subsection 9.1.2, we will illustrate the first of these two methods by a numerical example.

9.1. Exponential Demand Case

Let the demand densities be denoted as

$$g_0(z) = \delta_0 e^{-\delta_0 z}, \quad g_1(z) = \delta_1 e^{-\delta_1 z}, \quad \delta_0 > \delta_1. \tag{96}$$

Since their mean demands satisfy $1/\delta_0 < 1/\delta_1$, the densities represent the low and high demands environments, respectively. Finally, we have the functional equation

$$\begin{aligned} Z'(x,\beta) = \Big(&(c(1-\alpha)+\alpha h)(\beta+1) - \alpha(h+p)(\beta e^{-\delta_0 x^+} + e^{-\delta_1 x^+}) \\ &+ \alpha\int_0^{+\infty} \delta_1 e^{-\delta_1 z} Z'\Big(x-z, \beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big)dz\Big)^+. \end{aligned} \tag{97}$$

9.1.1. Approximation through Iteration

We go back to (97) and consider the iteration

$$Z'_{k+1}(x,\beta) = \Big((c(1-\alpha)+\alpha h)(\beta+1) - \alpha(h+p)(\beta e^{-\delta_0 x^+} + e^{-\delta_1 x^+}) \\ + \alpha \int_0^{+\infty} \delta_1 e^{-\delta_1 z} Z'_k\Big(x-z, \beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big)dz\Big)^+, \tag{98}$$

starting with

$$Z'_1(x,\beta) = \Big((c(1-\alpha)+\alpha h)(\beta+1) - \alpha(h+p)(\beta e^{-\delta_0 x^+} + e^{-\delta_1 x^+})\Big)^+. \tag{99}$$

We define first $S_1(\beta)$ by solving

$$\beta e^{-\delta_0 S_1(\beta)} + e^{-\delta_1 S_1(\beta)} = (\beta+1)\frac{c(1-\alpha)+\alpha h}{\alpha(h+p)}, \quad S_1(\beta) > 0. \tag{100}$$

The function $S_1(\beta)$ is decreasing and $Z'_1(x,\beta)$ is given by

$$Z'_1(x,\beta) = \begin{vmatrix} 0, \text{ if } x < S_1(\beta), \\ (c(1-\alpha)+\alpha h)(\beta+1) - \alpha(h+p)(\beta e^{-\delta_0 x} + e^{-\delta_1 x}), \text{ if } x \geq S_1(\beta). \end{vmatrix} \tag{101}$$

Note that a straightforward but lengthy calculation shows (see EC.8 for details) that $Z_2^{'}(x,\beta)$ can be obtained as

$$Z'_2(x,\beta) = \begin{vmatrix} 0, \quad \text{if} \quad x < S_2(\beta), \\ \\ (c(1-\alpha)+\alpha h)\big[\beta\big(1+\alpha(1-e^{-\delta_0 z_1(x,\beta)})\big) + 1 + \alpha(1-e^{-\delta_1 z_1(x,\beta)})\big] \\ -\alpha(h+p)\Big[e^{-\delta_0 x}\beta\big(1+\alpha\delta_0 z_1(x,\beta)\big) + e^{-\delta_1 x}\big(1+\alpha\delta_1 z_1(x,\beta)\big)\Big], \quad \text{if} \quad x \geq S_2(\beta), \end{vmatrix} \tag{102}$$

where $x = S_2(\beta) > 0$ is the solution of

$$(c(1-\alpha)+\alpha h)\big[\beta\big(1+\alpha(1-e^{-\delta_0 z_1(x,\beta)})\big) + 1 + \alpha(1-e^{-\delta_1 z_1(x,\beta)})\big] \\ -\alpha(h+p)\Big[e^{-\delta_0 x}\beta\big(1+\alpha\delta_0 z_1(x,\beta)\big) + e^{-\delta_1 x}\big(1+\alpha\delta_1 z_1(x,\beta)\big)\Big] = 0. \tag{103}$$

For $k \geq 2$, we can only consider approximate solutions for (98). First, we note that the upper limit for the integral in (98) can be replaced by x because $Z_k'(x,\beta) = 0$ for all β when $x \leq 0$. Let

$$R = \{(x,\beta) \mid 0 \leq x \leq X, \; 0 \leq \beta \leq B\} \tag{104}$$

be a rectangle in the $x\beta$-plane, and define a grid on R with grid spacings Δx and $\Delta\beta$ such that $N\Delta x = X$ and $M\Delta\beta = B$. We compute $Z_{k+1}'(x,\beta)$ at the grid points $(x_n, \beta_m) = (n\Delta x, m\Delta\beta)$, $0 \leq n \leq N$, $0 \leq m \leq M$, by approximating the integral in (98) using a composite trapezoidal formula as follows

$$\begin{aligned} Z_{k+1}'(x_n,\beta_m) = \Big(&(c(1-\alpha)+\alpha h)(\beta_m+1) - \alpha(h+p)(\beta_m e^{-2\gamma x_n^+} + e^{-\gamma x_n^+}) \\ &+ \alpha \frac{\Delta x}{2}\Big[\gamma Z_k'\big(x_n, 2\beta_m\big) + 2\sum_{l=1}^{N-1} \gamma e^{-\gamma z_l} Z_k'\big(x_n - z_l, 2\beta_m e^{-\gamma z_l}\big) \\ &+ \gamma e^{-\gamma x_n} Z_k'\big(0, 2\beta_m e^{-\gamma x_n}\big)\Big]\Big)^+, \end{aligned} \tag{105}$$

where $z_l = l\Delta x$, $l = 1,2,\cdots,N-1$. Note that since $Z_1'(x,\beta)$ is given explicitly by (99) during the computation of $Z_2'(x_n,\beta_m)$, $n = 0,1,2,\cdots,N$, $m = 0,1,\cdots,M$, the right hand side of (105) is known exactly. For $k \geq 2$, the second component of $\big(x_n - z_l, 2\beta_m e^{-\gamma z_l}\big)$ in general is not a grid value and, therefore, we obtain $Z_k'\big(x_n - z_l, 2\beta_m e^{-\gamma z_l}\big)$ by interpolation using the neighboring grid point values.

Now for illustrative purpose, we set the previous parameter values as in Section 8, and let $N = 40$, $M = 80$, with $\Delta x = 0.1$, $\Delta\beta = 0.05$. Below we show the structure of the curves $Z_1'(x,\beta)$, $Z_2'(x,\beta)$, and $Z_3'(x,\beta)$ in $\mathbb{R}^3$.

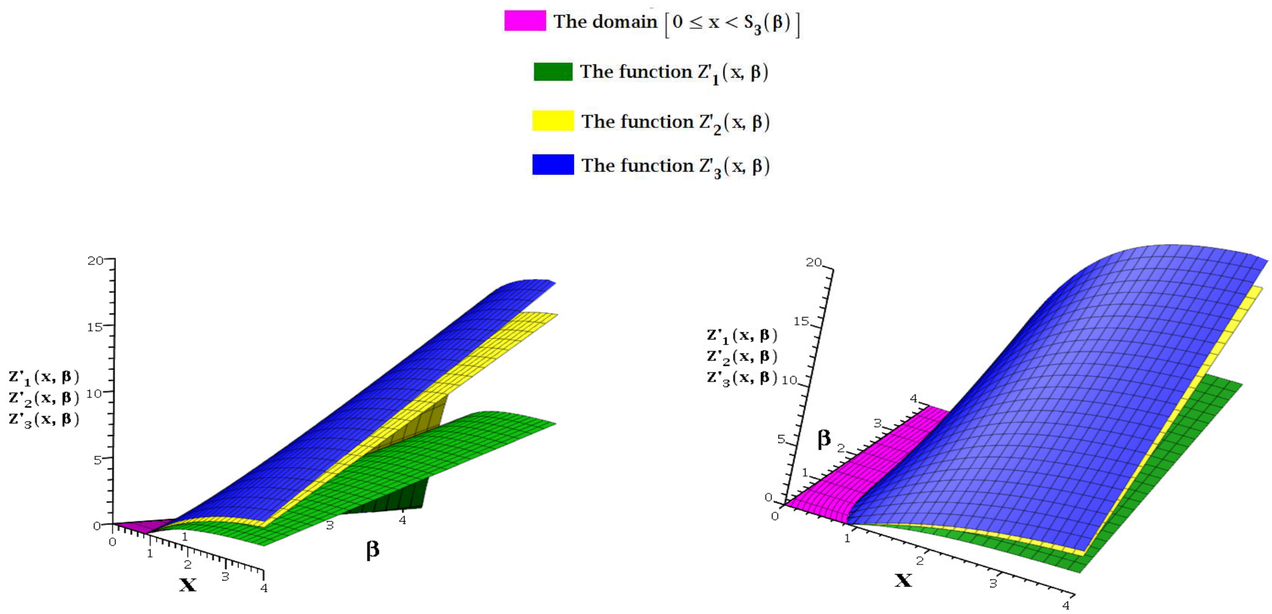

Figure 3 **The functions** $Z_1'(x,\beta), Z_2'(x,\beta), Z_3'(x,\beta)$.

Knowing the data points $\left(x_n, \beta_m, Z_k'(x_n,\beta_m)\right)_{\substack{n=0,\cdots,N\\ m=0,\cdots,M}}$, we can extract the values

$$S_k(\beta_m) = \sup\{x_n \mid Z_k'(x_n,\beta_m) = 0\}. \tag{106}$$

In Table 5 we display $S_k(\beta)$ for $k = 1,2,3,4$ and $\beta = 0,1,2,\cdots,10$.

Table 5 **Approximation of** $S_1(\beta)$, $S_2(\beta)$, $S_3(\beta)$, $S_4(\beta)$.

β	0	1	2	3	4	5	6	7	8	9	10
$S_1(\beta)$	0.8622	0.6048	0.5376	0.5076	0.4908	0.4800	0.4725	0.4670	0.4628	0.4594	0.4567
$S_2(\beta)$	0.8622	0.6021	0.5360	0.5067	0.4901	0.4795	0.4722	0.4667	0.4626	0.4593	0.4566
$S_3(\beta)$	0.8622	0.6017	0.5354	0.5059	0.4900	0.4795	0.4722	0.4667	0.4626	0.4593	0.4566
$S_4(\beta)$	0.8622	0.6012	0.5351	0.5058	0.4900	0.4795	0.4722	0.4667	0.4626	0.4593	0.4566
$S_3(\beta) - S_4(\beta)$	0	$5\ 10^{-4}$	$3\ 10^{-4}$	10^{-4}	0	0	0	0	0	0	0

Based on Table 5 we use the approximation $S(\beta) \approx S_4(\beta)$.

9.1.2. Computation of Base Stock for Exponential Demand

We now consider the demand densities (96) with $\delta_0 = 2$ and $\delta_1 = 1$, and the parameter values $\alpha = 0.9$, $p = 4$, $c = 10$, and $h = 1$ as in Section 8. To show how to implement our model, we suppose that the true demand density is $g_0(z) = e^{-\delta_0 z}$. Then the true mean is $1/\delta_0 = 1/2$ and the true base stock as computed using (6) is $S = 0.4311$. As in Subsection 8.3, the IM does not know the true demand, and so he must make an initial guess of the hyperparamaters $\beta_0 \geq 0$ and $\beta_1 > 0$ such that $\beta_0 + \beta_1 = 1$ and $\beta = \beta_0/\beta_1$ is defined.

We use the predefined Matlab function *exprnd*$(1/2, 1, 5)$ to generate $N = 5$ simulated demands according to the true probability density e^{-2z}. These are $D_1 = 0.1522$, $D_2 = 0.9613$, $D_3 = 0.1323$, $D_4 = 0.0338$, and $D_5 = 0.2135$.

In view of (92), we obtain the posteriors of β_0 and β_1 as follows:

$$posterior\ \ \beta_0 = \frac{\beta_0 g_0(D_1)}{\beta_0 g_0(D_1) + \beta_1 g_1(D_1)} \text{ and } posterior\ \ \beta_1 = \frac{\beta_1 g_1(D_1)}{\beta_0 g_0(D_1) + \beta_1 g_1(D_1)}. \tag{107}$$

It is straightforward to see from (107) that $posterior\ \ \beta_0 + posterior\ \ \beta_1 = 1$ in each period. We will illustrate our model with two different guesses: $(\beta_0, \beta_1) = (0.6, 0.4)$ and $(\beta_0, \beta_1) = (0.4, 0.6)$, respectively. These gives us the initial $\beta = 1.5$ and $\beta = 2/3$.

We first start with an initial guess $(\beta_0, \beta_1) = (0.6, 0.4)$ and consider the inventory evolution (14), with, say, the initial inventory $x_1 = 0$. The initial $\beta = \beta_0/\beta_1 = 1.5$, for which we compute the base stock $S(1.5) = 0.5512$. So we order up to it and obtain the order quantity $v_1 = 0.5512$, which gives $x_2 = x_1 + v_1 - D_1 = 0.3990$. In period 2, $\beta_0 = \dfrac{0.6 g_0(D_1)}{0.6 g_0(D_1) + 0.4 g_1(D_1)} = 0.7207$, $\beta_1 = \dfrac{0.4 g_1(D_1)}{0.6 g_0(D_1) + 0.4 g_1(D_1)} = 0.2796$, $\beta = \dfrac{0.6 g_0(D_1)}{0.4 g_1(D_1)} = 4.432$, and the base stock $S(2.5776) = 0.5072$. Thus, the order quantity $v_2 = 0.1082$ and $x_3 = -0.4541$. Following a similar process as in the second period, we present the inventory dynamics for the first five periods in Table 6 starting with $\beta = 1.5$. Similarly, we obtain Table 7 with a different initial guess $\beta = 2/3$.

Table 6 **Optimal path starting with $\beta = 1.5$.**

D_n	0.1522	0.9613	0.1323	0.0338	0.2135
β	1.5	2.5776	1.9704	3.4522	6.6752
$S(\beta)$	0.5512	0.5072	0.5271	0.4882	0.4588
v_n	0.5512	0.1082	0.9812	0.0934	0.0044
x_n	0	0.3990	-0.4541	0.3948	0.4544

Table 7 **Optimal path starting with $\beta = 2/3$.**

D_n	0.1522	0.9613	0.1323	0.0338	0.2135
β	0.6667	1.1451	0.8757	1.5344	2.9668
$S(\beta)$	0.5512	0.5779	0.6065	0.5491	0.5074
v_n	0.5512	0.1788	0.9899	0.0749	0
x_n	0	0.3990	-0.3835	0.4741	0.5152

From the fourth row in each table, we see that the base stock appears to be moving toward its true value of 0.4311.

We now generate $N = 1000$ simulated demands according to the exponential distribution $g_0(z)$ and run the simulation for N periods. If $N = 1000$, we compute $S\big(1.5 \cdot 2^{N-1} e^{-\sum_{i=1}^{N-1} D_i}\big) = S\big(1.5 \cdot 2^{999} e^{-501.2155}\big) = 0.4312$ and $S\big(2/3 \cdot 2^{N-1} e^{-\sum_{i=1}^{N-1} D_i}\big) = S\big(2/3 \cdot 2^{999} e^{-501.2155}\big) = 0.4312$. We see that both values are the same and very close to the true base stock of 0.4311. We can also realize that in period N, our method has lead to the value of the hyperparameter β_0 very close to its true value of 1.

$$posterior \;\; \beta_0 = \frac{0.6 \cdot 2^{N-1} e^{-2\sum_{i=1}^{N-1} D_i}}{0.6 \cdot 2^{N-1} e^{-2\sum_{i=1}^{N-1} D_i} + 0.4 \cdot e^{-\sum_{i=1}^{N-1} D_i}} = \frac{0.6 \cdot 2^{999} e^{-1002.4310}}{0.6 \cdot 2^{999} e^{-1002.4310} + 0.4 \cdot e^{-501.2155}}$$

$$= \frac{0.4 \cdot 2^{N-1} e^{-2\sum_{i=1}^{N-1} D_i}}{0.4 \cdot 2^{N-1} e^{-2\sum_{i=1}^{N-1} D_i} + 0.6 \cdot e^{-\sum_{i=1}^{N-1} D_i}} = \frac{0.4 \cdot 2^{999} e^{-1002.4310}}{0.4 \cdot 2^{999} e^{-1002.4310} + 0.6 \cdot e^{-501.2155}} \approx 1.$$

9.1.3. Approximation by Piecewise Constant Functions

Here we provide an alternate approximation procedure with technical details in EC.9 and show in EC.10 that it gives essentially the same results as obtained using the procedure in Subsection 9.1.1. Here we approximate the function $Z'(x, \beta)$ by a piecewise constant function in the hyperparameter β, so we write

$$Z'(x, \beta) \approx Z'_{[\beta]}(x), \quad [\beta] \le \beta < [\beta] + 1, \tag{108}$$

where $[\beta]$ is the integer part of β. We can therefore approximate (97) as

$$\begin{aligned} Z'_k(x) = \Big(& (c(1-\alpha)+\alpha h)(k+1) - \alpha(h+p)(ke^{-\delta_0 x^+} + e^{-\delta_1 x^+}) \\ & + \alpha\delta_1 \int_0^{+\infty} e^{-\delta_1 z} Z'\Big(x-z, k\frac{\delta_0}{\delta_1} e^{-(\delta_0-\delta_1)z}\Big) dz \Big)^+, \end{aligned} \tag{109}$$

where k is an integer. For instance, suppose $\delta_0/\delta_1 = 2$, then by setting $\delta_1 = \gamma$, we compute $Z'_k(x)$ by approximating the integral on its right hand side using the composite trapezoidal formula. See EC.9 for details.

We illustrate our findings by displaying $Z'_0(x)$, $Z'_1(x)$, $Z'_2(x)$, and $Z'_3(x)$ on Figure 4 for the special case of parameter values as in Subsection 9.1.1 and $\delta_1 = 1$.

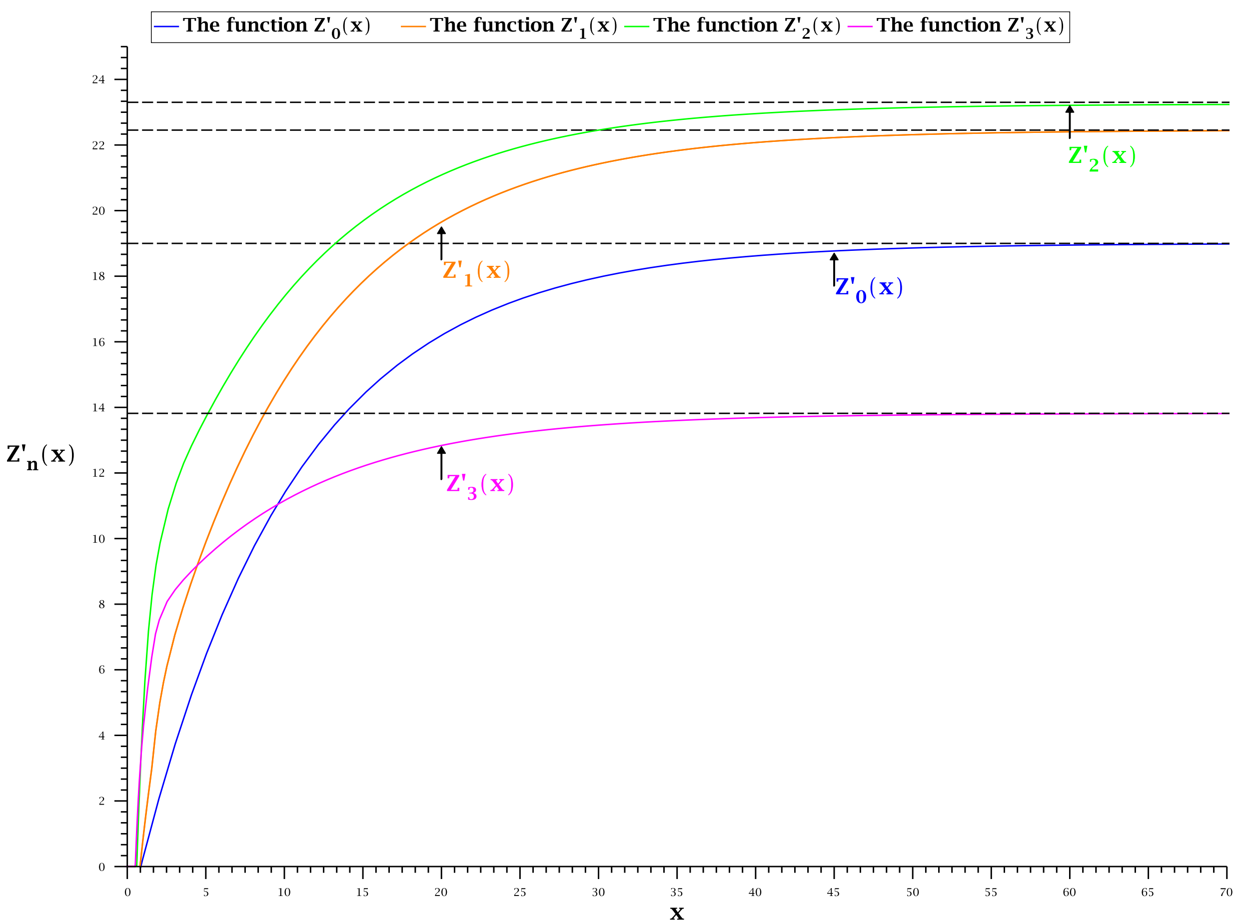

Figure 4 **The functions** $Z'_0(x)$, $Z'_1(x)$, $Z'_2(x)$, $Z'_3(x)$.

Since this method provides the approximation only at integer values of β, we show in EC.10 that it produces results that are very close to the method in Subsection 9.1.1 at the integer values of β.

10. Concluding Remarks

In this paper we have considered a standard discrete-time infinite-horizon inventory problem with an unknown general demand with backlog allowed. We begin with a general belief density prior and update it as the demand history unfolds. We analyze the resulting functional Bellman equation for the value function that depends on the current inventory level and the current (updated) infinite-dimensional belief. We prove the optimality of the base stock policy with the base stock dependent only on the current belief. We provide an algorithm to solve for the base stock policy and prove its convergence. We apply our methodology to the important case when the initial belief takes the form of a conjugate prior, which generalizes the related papers devoted to showing the optimality of the base stock policy in conjugate probability settings. Since we can set a completely general belief, we are also able to entertain the case when the true demand density belongs to a set of possible densities including nonparametric ones. To keep things simple, we only focus on the case of two possible densities. We illustrate our theory by solving a few numerical examples. Finally, given that our method is quite general, it is therefore applicable to inventory problems with lost sales, lead times, and fixed ordering costs. We leave analysis of these problems as future research topics.

References

Araman V, Caldentey R (2009) Dynamic pricing for nonperishable products with demand learning. *Oper. Res.* **57**(5) 1169-1188.

Aviv Y, Pazgal A (2005) A partially observed Markov decision process for dynamic pricing. *Management. Sci.* **51**(9) 1400-1416.

Azoury KS (1985) Bayes solution to dynamic inventory models under unknown demand distribution. *Management. Sci.* **31** 1150-1160.

Bensoussan A (2011) *Dynamic Programming and Inventory Control,* Studies in Probability, Optimization and Statistics, Vol. 3. (IOS Press, Amsterdam).

Bisi A, Dada M (2007) Dynamic learning, pricing, and ordering by a censored newsvendor. *Naval Res. Logist.* **54**(4) 448-461.

Cyganowski S, Kloeden P, Ombach J (2002) *From Elementary Probability to Stochastic Differential Equations with MAPLE®*. (Springer-Verlag, Berlin Heidelberg).

Dvoretzky A, Kiefer J, Wolfowitz J (1952) The inventory problem: II. Case of unknown distributions of demand. *Econometrica* **20** 450-466.

Farias V, Van Roy B, (2010) Dynamic pricing with a prior on market response. *Oper. Res.* **58**(1) 16-29.

Iglehart DL (1964) The dynamic inventory problem with unknown demand distribution. *Management. Sci.* **10** 429-440.

Karlin S (1960) Dynamic inventory policy with varying stochastic demands. *Management. Sci.* **6** 231-258.

Kay SM (2006) *Intuitive Probability and Random Processes using Matlab®*. (Springer, New York).

Larson CE, Olson LJ, Sharma S (2001) Optimal inventory policies when the demand distribution is not known. *J. Econom. Theory* **101** 281-300.

Lovejoy WS (1990) Myopic policies for some inventory models with uncertain demand distributions. *Management. Sci.* **36** 724-738.

Petruzzi NC, Dada M (2002) Dynamic pricing and inventory control with learning. *Naval Res. Logist.* **49** 303-325.

Porteus EL (1990) Stochastic inventory theory. Heyman DP, Sobel MJ, eds. *Handbooks in Operations Research and Management Science,* Vol. 2. (Elsevier, North-Holland, Netherlands), 605-652.

Roth MW (2018) *Modeling and Simulation of Everyday Things.* (CRC Press, Boca Raton, FLorida).

Scarf H (1959) Bayes solutions to the statistical inventory problem. *Ann. Math. Statist.* **30** 490-508.

Scarf H (1960a) Some remarks on Bayes solutions to the inventory problem. *Naval Res. Logist. Quart.* **7** 591-596.

Shingareva I, Celaya CL (2007) *Maple and Mathematica, A Problem Solving Approach for Mathematics.* Second edition, (Springer-Verlag, New York).

Subrahmanyan S, Shoemaker R (1996) Developing optimal pricing and inventory policies for retailers who face uncertain demand. *J. Retailing* **72** 7-30.

Wazwaz AM (2011) *Linear and Nonlinear Integral Equations, Methods and Applications.* (Higher Education Press Beijing and Springer-Verlag, Berlin Heidelberg).

Zhang J, Chen J (2006) Bayesian solution to pricing and inventory control under unknown demand distribution. *Oper. Res. Lett.* **34**(5) 517-524.

Proofs of Statements

EC.1. Brief Review of the Additional Literature on Partially Observed Inventory Control Problems with Demand Learning

Murray and Silver (1966) study a style good inventory problem of a firm where a number of known customers in a period purchase the firm's product with an unknown probability. Even though they assume unmet demands to be lost, they observe full demands and use the Beta prior to update the purchase probability. They do not characterize the optimal policy, but numerically solve a two-period example. On the other hand, Harpaz et al. (1982) take the Bayesian approach to the inventory problem with demand learning when unmet demands are lost and thus not fully observed. Braden and Freimer (1991) develop what they call newsvendor distributions that allow for a parsimonious updating process of the prior distribution in censored demand settings. Lariviere and Porteus (1999) consider an inventory problem of a new product that is about to be offered for sale under the lost sales regime. They use Weibull, a newsvendor distribution, to model market demand. They follow the methods of Scarf (1960a) and Azoury (1985) to obtain a dimensionality reduction. Miller (1986) shows that such reduction can also be achieved with exponentially smoothed forecasts. Under additional restrictions, Lariviere and Porteus (1999) obtain insights into the optimal order such as whether to drop a product or to overstock it to obtain more demand information. These restrictions ensure the distribution of demand in the future to be stochastically increasing in the demand that has been observed in the past. Hakzos and Seshadri (2004) named this property of distributions as conditional monotonicity and supplied the necessary and sufficient conditions for it to hold. Chen and Plambeck (2008) show that this "stock more" result can be reversed in the case of nonperishable inventory with unobserved lost sales. Bisi and Dada (2007) consider both the cases of perishable and nonperishable inventory under additive as well as multiplicative demand models.

Another stream of literature assumes general demand distributions and focus on showing that the conditional probability of demand and or inventory given past observations is a sufficient statistic, which in general is infinite dimensional. This allows the use of dynamic programming to analyze the problem and prove the existence and uniqueness of a solution of the resulting functional infinite-dimensional Bellman equation as well as the existence of an optimal feedback policy and some of its properties. This extends the idea of Kalman filtering to inventory problems, which are nonlinear in general, as carried out in (2016). The interested reader can consult Ding et al. (2002), Lu et al. (2005), Treharne and Sox (2002), Sethi (2010), Bensoussan et al.((2007), (January 2008), (February 2008), (2009a), (2009c), (2010), (2016)), and references therein.

(EC.1)

EC.2. Proof of Lemma 1

For the lower bound, let the control sequence $V=\{y_1,\cdots,y_n,\cdots\}$, with $y_1=y\geq x$, $y_n\geq x_n$. Then $-cx_n+cy_n\geq 0$. We first obtain estimates. We have

$$x_n \geq x_{n-1}-D_{n-1},$$

hence,

$$x_n^+ \geq x_{n-1}^+ - D_{n-1} \geq x_{n-2}^+ - D_{n-1} - D_{n-2} \geq \cdots \geq x^+ - \sum_{i=1}^{n-1} D_i.$$

Next, we have

$$\begin{aligned}
J_{x,f(.)}(V) &= \sum_{n=1}^{+\infty} \alpha^{n-1} E\big[hx_n^+ + px_n^- + \underbrace{c(y_n-x_n)}_{\geq 0}\big], \\
&\geq \sum_{n=1}^{+\infty} \alpha^{n-1} E\big[hx_n^+ + px_n^-\big] \\
&\geq \sum_{n=1}^{\infty} \alpha^{n-1} E\big[hx^+ - \sum_{i=1}^{n-1} D_i\big] + px^- \\
&= \frac{hx^+}{1-\alpha} + px^- - h\sum_{n=1}^{\infty}\alpha^{n-1}E[(n-1)D] \\
&= \frac{hx^+}{1-\alpha} + px^- - hE[D]\sum_{n=1}^{\infty}\alpha^{n-1}(n-1) \\
&= \frac{hx^+}{1-\alpha} + px^- - \frac{h\alpha}{(1-\alpha)^2}E[D],
\end{aligned} \tag{EC.2}$$

where the arithmetico-geometric series

$$\sum_{n=1}^{\infty}\alpha^{n-1}(n-1)=\sum_{n=0}^{\infty}n\alpha^n=\frac{\alpha}{(1-\alpha)^2},$$

and where we assume that

$$E[D]=\int E[D|\eta]f(\eta)d\eta=\int\varphi(\eta)f(\eta)d\eta<+\infty,\ \text{ with }\ \varphi(\eta)=\int_0^{+\infty}\bar{G}(z|\eta)dz.$$

Hence,

$$\Phi(x,f(.))=\inf_V J_{x,f(.)}(V)\geq\frac{hx^+}{1-\alpha}+px^- -\frac{h\alpha}{(1-\alpha)^2}\int\varphi(\eta)f(\eta)d\eta. \tag{EC.3}$$

Of course, this estimate is not useful when $|x|$ is small, since we already know that $\Phi(x,f(.))$ is positive. However, it will be for large values of $|x|$. We can thus replace the right hand side of equation (EC.3) by its positive part, and write

$$\Phi(x,f(.))\geq\Big(\frac{hx^+}{1-\alpha}+px^- -\frac{h\alpha}{(1-\alpha)^2}\int\varphi(\eta)f(\eta)d\eta\Big)^+. \tag{EC.4}$$

We next obtain an estimates from above by evaluating the cost of a specific policy. We can take in particular the control $y_n=x_n^+$. Denoting this control sequence by V^+, we get

$$J_{x,f(.)}(V^+)=\sum_{n=1}^{+\infty}\alpha^{n-1}E\big[hx_n^+ +(p+c)x_n^-\big].$$

We have $x_{n+1}^-\leq D_n$ and $x_{n+1}=x_n^+-D_n\leq x_n^+$, so $x_n\leq x_{n-1}^+$ and $x_{n+1}^+\leq x_n^+\leq\cdots\leq x^+$. Then,

$$J_{x,f(.)}(V^+)\leq\frac{hx^+}{1-\alpha}+(p+c)x^- +(p+c)\sum_{n=2}^{\infty}\alpha^{n-1}E[D].$$

Hence,

$$\Phi(x,f(.))=\inf_V J_{x,f(.)}(V)\leq\frac{hx^+}{1-\alpha}+(p+c)x^- +\frac{\alpha(p+c)}{1-\alpha}\int\varphi(\eta)f(\eta)d\eta, \tag{EC.5}$$

and, of course, the right hand side is a real constraint since we have assumed

$$\int\varphi(\eta)f(\eta)d\eta<+\infty. \tag{EC.6}$$

□

EC.3. Proof of Proposition 1

We shall prove that

$$\Phi(x,f(.)) \le \Phi^*(x,f(.)) = J_{x,f(.)}(V),\ \ \forall V. \tag{EC.7}$$

If $\Phi(x,f(.))$ solves (29) with the inequalities (23), then for $y \ge x$ we can assert that

$$\Phi(x,f(.)) \le hx^+ + px^- - cx + cy + \alpha E\Phi\Big(y-D, \frac{f(.)g(D|.)}{\int f(\eta)g(D|\eta)d\eta}\Big). \tag{EC.8}$$

Take now any $V=(y_1,\cdots,y_n,\cdots)$ and by (18) and (19), we construct sequentially the state (x_n, $f_n(.)$), with $y_n \ge x_n$. We can apply (EC.8) with $x = x_n$ and $f(.) = f_n(.)$. Note that

$$E\int \varphi(\eta)f_n(\eta)d\eta = \int \varphi(\eta)f(\eta)d\eta < +\infty,$$

therefore, $\int \varphi(\eta)f_n(\eta)d\eta < +\infty$, a.s. Taking the account of (18) and (19) and by a clear interpretation, we can write

$$\begin{aligned}\Phi(x_n,f_n(.)) &\le hx_n^+ + px_n^- - cx_n + cy_n + \alpha E\Phi\Big(y_n - D_n, \frac{f_n(.)g(D_n|.)}{\int f_n(\eta)g(D_n|\eta)d\eta}\Big)\\ &= hx_n^+ + px_n^- + c(y_n-x_n) + \alpha E\big[\Phi(x_{n+1},f_{n+1}(.))|\mathcal{D}^{n-1}\big].\end{aligned} \tag{EC.9}$$

Therefore,

$$\alpha^{n-1}E\Phi(x_n,f_n(.)) \le \alpha^{n-1}E\big[hx_n^+ + px_n^- + c(y_n-x_n)\big] + \alpha^n E\Phi(x_{n+1},f_{n+1}(.)). \tag{EC.10}$$

Summing these relations between $n=1$ and $n=N$, it follows

$$\Phi(x,f(.)) \le \sum_{n=1}^{N} \alpha^{n-1}E\big[hx_n^+ + px_n^- + c(y_n-x_n)\big] + \alpha^N E\Phi(x_{N+1},f_{N+1}(.)). \tag{EC.11}$$

But from the right inequality of (EC.5), we get

$$E\Phi(x_{N+1},f_{N+1}(.)) \le \frac{h}{1-\alpha}Ex_{N+1}^+ + (p+c)Ex_{N+1}^- + \frac{\alpha(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta.$$

Thanks to the admissibility interval $\mathcal{I}_{x,f(.)}$, we have

$$\left|\begin{aligned} &x_{N+1}^+ \le x^+ + \frac{1}{c}\frac{\alpha N}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\int \varphi(\eta)f(\eta)d\eta,\\ &\\ &x_{N+1}^- \le x^- + D_1 + \cdots + D_N.\end{aligned}\right. \tag{EC.12}$$

It follows immediately that $\alpha^N E\Phi(x_{N+1}, f_{N+1}(.)) \to 0$ as $N \to +\infty$. Therefore, from (EC.11) we obtain $\Phi(x, f(.)) \leq J_{x,f(.)}(V)$, from which we get the result (EC.7). $\square$

EC.4. Proof of Theorem 1

The proof is long and it is divided into four segments or subsections.

Preliminaries

From (40), we neglect the sharper constraint on the left of $W(x, f(.))$, introduce the set Γ of functionals on $\mathbb{R} \times L^1_{\varphi+}(\mathbb{R})$ such that

$$\Gamma = \Bigg\{ W(x, f(.)) | \ W \text{ is continuous}, \\ 0 \leq W(x, f(.)) \leq \Big(\frac{hx^+}{1-\alpha} + (p+c)x^- \Big) \int f(\eta) d\eta + \frac{\alpha(p+c)}{1-\alpha} \int \varphi(\eta) f(\eta) d\eta \Bigg\}, \tag{EC.13}$$

and define the operator on Γ as follows:

$$T(W)(x, f(.)) = (hx^+ + px^- - cx) \int f(\eta) d\eta + \inf_{y \in \mathcal{L}_{x,f(.)}} \big[cy \int f(\eta) d\eta + \alpha \int_0^{+\infty} W(y - z, f(.)g(z|.)) dz \big]. \tag{EC.14}$$

This operator is monotone in the following sense:

$$W(x, f(.)) \leq \tilde{W}(x, f(.)), \ \forall x, f(.) \Longrightarrow T(W)(x, f(.)) \leq T(\tilde{W})(x, f(.)), \ \forall x, f(.). \tag{EC.15}$$

The operator T maps Γ into itself

We begin with the following important result.

Proposition EC.1. *Assume the properties (31)-(33). The operator T maps Γ into itself.*

Proof. From the assumptions, we obtain immediately that $f(.)g(z|.)$ belongs to $L^{1+}_{\varphi}(\mathbb{R})$ for any $z > 0$. Since W belongs to Γ, for any $z > 0$, $W(y - z, f(.)g(z|.))$ is well defined and

$$0 \leq W(y-z, f(.)g(z|.)) \leq \Big(\frac{h(y-z)^+}{1-\alpha} + (p+c)(y-z)^- \Big) \int f(\eta) g(z|\eta) d\eta + \frac{\alpha(p+c)}{1-\alpha} \int \varphi(\eta) f(\eta) g(z|\eta) d\eta.$$

This is well defined since $f(.)g(z|.)$ belongs to $L^1_{\varphi+}(\mathbb{R})$. We may integrate in $z>0$ to obtain

$$\int_0^{+\infty} W(y-z, f(.)g(z|.))dz \leq \int_0^{+\infty} \Big(\frac{h(y-z)^+}{1-\alpha} + (p+c)(y-z)^-\Big) \int f(\eta)g(z|\eta)d\eta dz + \frac{\alpha(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta.$$

Therefore,

$$\begin{aligned} T(W)(x,f(.)) &\leq (hx^+ + px^- - cx)\int f(\eta)d\eta + \frac{\alpha^2(p+c)}{1-\alpha}\int \varphi(\eta)f(\eta)d\eta \\ &+ \inf_{y\in\mathcal{L}_{x,f(.)}} \left(cy\int f(\eta)d\eta + \alpha\int_0^{+\infty}\Big(\frac{h(y-z)^+}{1-\alpha} + (p+c)(y-z)^-\Big)\int f(\eta)g(z|\eta)d\eta dz\right). \end{aligned} \tag{EC.16}$$

We bound the right hand side of (EC.16) by picking $y=x^+$, and use $(x^+-z)^+\leq x^+$, $(x^+-z)^-\leq z$, and obtain

$$\begin{aligned} &T(W)(x,f(.)) \\ &\leq (hx^+ + (p+c)x^-)\int f(\eta)d\eta + \frac{\alpha^2(p+c)}{1-\alpha}\int\varphi(\eta)f(\eta)d\eta + \frac{\alpha h x^+}{1-\alpha}\int f(\eta)d\eta + \alpha(p+c)\int\varphi(\eta)f(\eta)d\eta \\ &= \Big(\frac{hx^+}{1-\alpha} + (p+c)x^-\Big)\int f(\eta)d\eta + \frac{\alpha(p+c)}{1-\alpha}\int\varphi(\eta)f(\eta)d\eta. \end{aligned}$$

Since, obviously $T(W)(x,f(.))\geq 0$, the constraint in (EC.13) is satisfied for $T(W)(x,f(.))$. We have next to prove that the map $x,f(.)\to T(W)(x,f(.))$ is continuous. We first define for $y\geq x$,

$$J(W)(x,f(.))(y) = cy\int f(\eta)d\eta + \alpha\int_0^{+\infty} W(y-z, f(.)g(z|.))dz. \tag{EC.17}$$

Since $W(x,f(.))$ is continuous in both arguments, the function $y\to J(W)(x,f(.))(y)$ is continuous and thus attains its minimum in y on the bounded interval $\mathcal{L}_{x,f(.)}$. Then, by the measurable selection theorem (see, e.g., Bensoussan et al. (1983) or Bensoussan (2011)), we can find a Borel function $\hat{y}(x,f(.))$ that realizes the infimum in (EC.14) for any $(x,f(.))$ in $\mathbb{R}\times L^1_{\varphi+}(\mathbb{R})$ i.e.,

$$\left|\begin{array}{l} T(W)(x,f(.)) = (hx^+ + px^- - cx)\int f(\eta)d\eta + J(W)(x,f(.))(\hat{y}(x,f(.))), \ \ \forall x, f(.), \\ \\ \hat{y}(x,f(.))\in\mathcal{L}_{x,f(.)}. \end{array}\right.$$

Let us show that $x, f(.) \to T(W)(x, f(.))$ is continuous. Let $x_n,\ x \in \mathbb{R},\ x_n \to x$ and $f_n(.),\ f(.) \in L^1_{\varphi+}(\mathbb{R}),\ f_n(.) \to f(.)$ in $L^1_{\varphi+}(\mathbb{R})$. We call $y_n = \hat{y}(x_n, f_n(.))$. Clearly y_n is bounded, so we can extract a subsequence, still denoted $x_n,\ y_n,\ f_n(.) \to x,\ \hat{y},\ f(.)$. Since

$$x_n \le y_n \le x_n^+ + \frac{1}{c}\frac{\alpha}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\frac{\int \varphi(\eta) f_n(\eta) d\eta}{\int f_n(\eta) d\eta},$$

we get

$$x \le \hat{y} \le x^+ + \frac{1}{c}\frac{\alpha}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\frac{\int \varphi(\eta) f(\eta) d\eta}{\int f(\eta) d\eta}.$$

From the assumptions (31)-(33), we can assert that

$$f_n(.) g(z|.) \to f(.) g(z|.) \text{ in } L^1_{\varphi+}(\mathbb{R}),\ \ \forall z > 0.$$

From the continuity of W, it follows that

$$W(y_n - z, f_n(.) g(z|.)) \to W(y - z, f(.) g(z|.)),\ \ \forall z > 0.$$

Moreover,

$$W(y_n - z, f_n(.) g(z|.)) \le \Big(\frac{h(y_n - z)^+}{1-\alpha} + (p+c)(y_n - z)^-\Big) \int f_n(\eta) g(z|\eta) d\eta \\ + \frac{\alpha(p+c)}{1-\alpha} \int \varphi(\eta) f_n(\eta) g(z|\eta) d\eta.$$

Since $f_n(.) \to f(.)$ in $L^1_{\varphi+}(\mathbb{R})$, we can find a new subsequence, still denoted by $f_n(.)$ such that

$$\int (\varphi(\eta) + 1) \sup_n f_n(\eta) d\eta < +\infty.$$

This is a classical result whose proof is as follows. The sequence $\varphi(\eta) f_n(\eta)$ is a Cauchy sequence in $L^1(\mathbb{R})$. Therefore, there exists a subsequence $\varphi(\eta) f_{n_j}(\eta)$ such that

$$\int (\varphi(\eta) + 1) |f_{n_{j+1}}(\eta) - f_{n_j}(\eta)| d\eta \le \frac{1}{2^j}.$$

Then,

$$h(\eta) \overset{\text{def}}{=} (\varphi(\eta) + 1) \Big[f_{n_o}(\eta) + \sum_{j=0}^{+\infty} |f_{n_{j+1}}(\eta) - f_{n_j}(\eta)| \Big]$$

is in $L^1(\mathbb{R})$ and $(\varphi(\eta)+1)f_{n_j}(\eta) \le h(\eta)$, which implies

$$\int (\varphi(\eta)+1) \sup_j f_{n_j}(\eta) d\eta < +\infty.$$

So, for some subsequence, and using (31), we can assert that for $z>0$,

$$\begin{aligned} &W(y_n - z, f_n(.)g(z|.)) \\ &\le \Big(\frac{h(y_n-z)^+}{1-\alpha} + (p+c)(y_n-z)^-\Big) \int \sup_n f_n(\eta) g(z|\eta) d\eta + \frac{\alpha(p+c)}{1-\alpha} \int \varphi(\eta) \sup_n f_n(\eta) g(z|\eta) d\eta \\ &\le \Big(\frac{h \sup_n y_n^+}{1-\alpha} + (p+c)(\sup_n y_n^- + z)\Big) \int \sup_n f_n(\eta) g(z|\eta) d\eta + \frac{\alpha(p+c)}{1-\alpha} \int \varphi(\eta) \sup_n f_n(\eta) g(z|\eta) d\eta, \end{aligned} \tag{EC.18}$$

which is a fixed integrable function of z. So, we can apply Lebesgue's dominated convergence theorem (see, e.g., Royden (1968) or Brezis (2011)) to assert that

$$\int_0^{+\infty} W(y_n - z, f_n(.)g(z|.)) dz \to \int_0^{+\infty} W(\hat{y} - z, f(.)g(z|.)) dz. \tag{EC.19}$$

Going back to (EC.17), we obtain

$$\begin{aligned} T(W)(x_n, f_n(.)) &= (hx_n^+ + px_n^- - cx_n) \int f_n(\eta) d\eta + J(W)(x_n, f_n(.))(y_n) \\ &\to (hx^+ + (p+c)x^-) \int f(\eta) d\eta + J(W)(x, f(.))(\hat{y}). \end{aligned} \tag{EC.20}$$

It remains to show that

$$(hx^+ + (p+c)x^-) \int f(\eta) d\eta + J(W)(x, f(.))(\hat{y}) = T(W)(x, f(.)). \tag{EC.21}$$

Take first $y \in \Big(x, x^+ + \dfrac{1}{c} \dfrac{\alpha}{(1-\alpha)^2} (h + (p+c)(1-\alpha)) \dfrac{\int \varphi(\eta) f(\eta) d\eta}{\int f(\eta) d\eta}\Big)$. For n large enough we have

$$x_n < y < x_n^+ + \frac{1}{c} \frac{\alpha}{(1-\alpha)^2} (h + (p+c)(1-\alpha)) \frac{\int \varphi(\eta) f_n(\eta) d\eta}{\int f_n(\eta) d\eta},$$

and thus,

$$T(W)(x_n, f_n(.)) \ge (hx_n^+ + (p+c)x_n^-) \int f_n(\eta) d\eta + J(W)(x_n, f_n(.))(y).$$

Going to the limit, this implies that

$$(hx^+ + (p+c)x^-) \int f(\eta) d\eta + J(W)(x, f(.))(\hat{y}) \ge (hx^+ + (p+c)x^-) \int f(\eta) d\eta + J(W)(x, f(.))(y).$$

Also,

$$T(W)(x_n, f_n(.)) \geq (hx_n^+ + (p+c)x_n^-) \int f_n(\eta)d\eta + J(W)(x_n, f_n(.))(x_n),$$

$$T(W)(x_n, f_n(.)) \geq (hx_n^+ + (p+c)x_n^-) \int f_n(\eta)d\eta$$
$$+ J(W)(x_n, f_n(.))\Big(x_n^+ + \frac{1}{c}\frac{\alpha}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\frac{\int \varphi(\eta) f_n(\eta)d\eta}{\int f_n(\eta)d\eta}\Big),$$

and going to the limit, it follows that

$$(hx^+ + (p+c)x^-) \int f(\eta)d\eta + J(W)(x, f(.))(\hat{y}) \geq (hx^+ + (p+c)x^-) \int f(\eta)d\eta + J(W)(x, f(.))(x),$$

$$(hx^+ + (p+c)x^-) \int f(\eta)d\eta + J(W)(x, f(.))(\hat{y}) \geq (hx^+ + (p+c)x^-) \int f(\eta)d\eta$$
$$+ J(W)(x, f(.))\Big(x^+ + \frac{1}{c}\frac{\alpha}{(1-\alpha)^2}(h+(p+c)(1-\alpha))\frac{\int \varphi(\eta) f(\eta)d\eta}{\int f(\eta)d\eta}\Big).$$

Collecting results, we obtain

$$(hx^+ + (p+c)x^-) \int f(\eta)d\eta + J(W)(x, f(.))(\hat{y}) \geq T(W)(x, f(.)).$$

Since the opposite is obviously true, we obtain (EC.21). So we have proven that $T(W)(x_n, f_n(.)) \rightarrow T(W)(x, f(.))$, at least for the subsequence that has been constructed. But for any converging subsequence, a similar reasoning will prove that the limit is $T(W)(x, f(.))$. This implies that the whole sequence $T(W)(x_n, f_n(.))$ converges towards $T(W)(x, f(.))$, which proves the continuity and completes the proof of the proposition. $\square$

Existence of a Solution

We proceed with the proof of existence of a solution of (39) and (40). In fact (39) is a fixed point equation. The functional $W(x, f(.))$ must satisfy

$$W(x, f(.)) = T(W)(x, f(.)). \tag{EC.22}$$

We define two iterations

$$\left| \begin{array}{l} W_{n+1}(x, f(.)) = T(W_n)(x, f(.)), \\ \\ W_0(x, f(.)) = 0. \end{array} \right. \tag{EC.23}$$

$$\left|\begin{array}{l} W^{n+1}(x,f(.)) = T(W^n)(x,f(.)), \\ W^0(x,f(.)) = (\dfrac{hx^+}{1-\alpha} + (p+c)x^-)\displaystyle\int f(\eta)d\eta + \dfrac{\alpha(p+c)}{1-\alpha}\displaystyle\int \varphi(\eta)f(\eta)d\eta. \end{array}\right. \tag{EC.24}$$

It is standard to check the sequence of inequalities

$$W_0 \le W_1 \le \cdots W_n \le \cdots \le W^0, \tag{EC.25}$$

$$W^0 \ge W^1 \ge \cdots W^n \ge \cdots \ge W_0. \tag{EC.26}$$

The first iteration is called the monotone increasing iteration and the second one is called the monotone decreasing iteration. It follows that

$$W_n(x,f(.)) \uparrow \underline{W}(x,f(.)) \text{ pointwise, } W^n(x,f(.)) \downarrow \overline{W}(x,f(.)) \text{ pointwise,} \tag{EC.27}$$

and necessarily,

$$W_0 \le \underline{W} \le \overline{W} \le W^0. \tag{EC.28}$$

We are going to show that $\underline{W}$ and $\overline{W}$ are solutions. First, from (EC.23) and the monotonicity property, we get

$$W_{n+1}(x,f(.)) \le T(\underline{W})(x,f(.)).$$

Going to the limit, we obtain

$$\underline{W}(x,f(.)) \le T(\underline{W})(x,f(.)), \ \ \forall x, f(.), \tag{EC.29}$$

and similarly,

$$\overline{W}(x,f(.)) \ge T(\overline{W})(x,f(.)), \ \ \forall x, f(.). \tag{EC.30}$$

The functions $W_n(x,f(.))$ and $W^n(x,f(.))$ are continuous. The limits are not necessarily continuous, but it is standard that

$$\underline{W}(x,f(.)) \text{ is l.s.c., } \quad \overline{W}(x,f(.)) \text{ is u.s.c.} \tag{EC.31}$$

Next, from (EC.23) again we have

$$W^{n+1}(x,f(.)) \le (hx^+ + px^- - cx)\int f(\eta)d\eta + J(W^n)(x,f(.))(y), \ \ \forall y \in \mathcal{L}_{x,f(.)}.$$

Looking at formula (EC.17) for $J(W^n)(x,f(.))(y)$, we have

$$W^n(y-z,\, f(.)g(z|.)) \downarrow \overline{W}(y-z,\, f(.)g(z|.)). \tag{EC.32}$$

Therefore, by Lebesgue's Lemma (see, e.g., Royden (1968) or Brezis (2011)) we have

$$\int_0^{+\infty} W^n(y-z, f(.)g(z|.))dz \downarrow \int_0^{+\infty} \overline{W}(y-z, f(.)g(z|.))dz, \tag{EC.33}$$

therefore, it follows easily that $J(W^n)(x,f(.))(y) \downarrow J(\overline{W})(x,f(.))(y)$. Also,

$$\overline{W}(x,f(.)) \le (hx^+ + px^- - cx)\int f(\eta)d\eta + J(\overline{W})(x,f(.))(y).$$

This implies $\overline{W}(x,f(.)) \le T(\overline{W})(x,f(.))$, since y is arbitrary in the interval $\mathcal{L}_{x,f(.)}$. From the opposite inequality (EC.30), we conclude that $\overline{W}(x,f(.))$ is a solution of (EC.22). Consider next the monotone increasing sequence $W_n(x,f(.))$ and $m > n$, so $W_m(x,f(.)) \ge W_n(x,f(.))$. We have

$$W_{m+1}(x,f(.)) = T(W_m)(x.f(.)) = (hx^+ + px^- - cx)\int f(\eta)d\eta + J(W_m)(x,f(.))(y_m)$$

since the inf is attained, in the definition of $T(W_m)(x.f(.))$, thanks to the continuity of $W_m(x,f(.))$, at some y_m, depending on $(x,f(.))$. Therefore,

$$W_{m+1}(x,f(.)) \ge (hx^+ + px^- - cx)\int f(\eta)d\eta + J(W_n)(x,f(.))(y_m).$$

Recalling that the sequence y_m is bounded, we can extract a subsequence that converges as $m \to +\infty$ to some y_*. We can pass to the limit in m in the preceding inequality to obtain

$$\underline{W}(x,f(.)) \ge (hx^+ + px^- - cx)\int f(\eta)d\eta + J(W_n)(x,f(.))(y_*).$$

Letting then $n \to +\infty$, we obtain

$$\underline{W}(x, f(.)) \geq (hx^+ + px^- - cx) \int f(\eta) d\eta + J(\underline{W})(x, f(.))(y_*) \geq T(\underline{W})(x, f(.)),$$

since y_* lies in the admissibility interval. Comparing with (EC.29) we see that $\underline{W}(x, f(.))$ is also a solution. So, we have two solutions of (EC.22), $\underline{W}(x, f(.))$ and $\overline{W}(x, f(.))$ with $\underline{W}(x, f(.)) \leq \overline{W}(x, f(.))$. The first one is l.s.c. and the second one is u.s.c. In fact $\underline{W}(x, f(.))$ is the smallest solution and $\overline{W}(x, f(.))$ is the largest solution, in the sense that if $W(x, f(.))$ is a solution satisfying (40), then necessarily

$$\underline{W}(x, f(.)) \leq W(x, f(.)) \leq \overline{W}(x, f(.)). \tag{EC.34}$$

Indeed, we have, by virtue of (40)

$$W_0(x, f(.)) \leq W(x, f(.)) \leq W^0(x, f(.)).$$

Suppose that for some n

$$W_n(x, f(.)) \leq W(x, f(.)) \leq W^n(x, f(.)). \tag{EC.35}$$

Then, by the monotonicity of the operator T

$$W_{n+1}(x, f(.)) = T(W_n)(x, f(.)) \leq T(W)(x, f(.)) = W(x, f(.)) \leq T(W^n)(x, f(.)) = W^{n+1}(x, f(.)),$$

so, (EC.35) holds for any n. Going to the limit, we obtain (EC.34). We check next that there exists a Borel function $\hat{y}(x, f(.))$ such that

$$\underline{W}(x, f(.)) = (hx^+ + px^- - cx) \int f(\eta) d\eta + J(\underline{W})(x, f(.))(\hat{y}(x, f(.))). \tag{EC.36}$$

This is because $\underline{W}(x, f(.))$ is l.s.c. We have to prove that

$$T(\underline{W})(x, f(.)) = (hx^+ + px^- - cx) \int f(\eta) d\eta + J(\underline{W})(x, f(.))(\hat{y}(x, f(.))).$$

But because $\underline{W}(x, f(.))$ is l.s.c., we see easily that $y \to J(\underline{W})(x, f(.))(y)$ is l.s.c. Therefore, we can find a Borel selection $\hat{y}(x, f(.))$ that attains the inf in y in the admissibility interval $\mathcal{L}_{x,f(.)}$.

Uniqueness of the Solution

Going back to functions $f(.)$ that are probability densities, we can transpose the results in W into results in Φ and thus we obtain that there is a minimum and a maximum solution $\underline{\Phi}(x,f(.))$ and $\overline{\Phi}(x,f(.))$ of (23) and (29). Recall that (29) is a rewriting of (22). Thanks to the feedback $\hat{y}(x,f(.))$, we construct the process $\hat{x}_n$, $\hat{y}_n$, and $\hat{f}_n(.)$ by the iterations

$$\hat{y}_n = \hat{y}(\hat{x}_n, \hat{f}_n(.)), \tag{EC.37}$$

$$\hat{x}_{n+1} = \hat{y}_n - D_n, \tag{EC.38}$$

$$\hat{f}_{n+1}(\xi) = \hat{f}_n(\xi)\frac{g(D_n|\xi)}{\int g(D_n|\eta)\hat{f}_n(\eta)d\eta}, \tag{EC.39}$$

starting with $\hat{x}_1 = x$, and $\hat{f}_1(.) = f(.)$. The measurability property of the selection $\hat{y}$ implies that the process $\hat{y}_n \geq \hat{x}_n$ is adapted to $\mathcal{D}^{n-1}$. We set $\hat{V} = (\hat{y}_1, \cdots, \hat{y}_n, \cdots)$ and define $J_{x,f(.)}(\hat{V})$. We are going to show that

$$\underline{\Phi}(x,f(.)) = J_{x,f(.)}(\hat{V}). \tag{EC.40}$$

Indeed, we proceed as in Proposition 1. We write the equation for $\underline{\Phi}(x,f(.))$ as

$$\underline{\Phi}(x,f(.)) = hx^+ + px^- - cx + c\hat{y}(x,f(.)) + \alpha E\underline{\Phi}\Big(\hat{y}(x,f(.)) - D, \frac{f(.)g(D|.)}{\int f(\eta)g(D|\eta)d\eta}\Big). \tag{EC.41}$$

Applying this relation with $x = \hat{x}_n,\ f(.) = \hat{f}_n(.)$, we can write as Proposition 1

$$\underline{\Phi}(\hat{x}_n, \hat{f}_n(.)) = h\hat{x}_n^+ + p\hat{x}_n^- - c\hat{x}_n + c\hat{y}_n + \alpha E[\underline{\Phi}(\hat{x}_{n+1}, \hat{f}_{n+1}(.))|\mathcal{D}^{n-1}].$$

Multiplying by α^{n-1}, taking the mathematical expectation and summing up for n running from 1 to N, we obtain, after elimination of identical terms in both sides,

$$\underline{\Phi}(x,f(.)) = \sum_{n=1}^{N} \alpha^{n-1} E[h\hat{x}_n^+ + p\hat{x}_n^- + c(\hat{y}_n - \hat{x}_n)] + \alpha^N E\underline{\Phi}(\hat{x}_{N+1}, \hat{f}_{N+1}(.)). \tag{EC.42}$$

Since $\underline{\Phi}(x,f(.)) \geq 0$, by letting $N \to +\infty$,

$$\underline{\Phi}(x,f(.)) \geq J_{x,f(.)}(\hat{V}) \geq \Phi^*(x,f(.)) = \inf_V J_{x,f(.)}(V).$$

Since $\underline{\Phi}(x, f(.))$ is the smallest solution satisfying (23), it follows that all solutions are larger than the value function. On the other hand, by the comparison result of Proposition 1, all solutions are smaller than the value function. Necessarily, the solution is unique and coincides with the value function. The feedback $\hat{y}(x, f(.))$ allows to construct an optimal control, by formulas (EC.37), (EC.38) and (EC.39). The base stock result is proved in Section 7. The proof of Theorem 1 has been completed. $\square$

EC.5. Proof of Lemma 2

From the relation (47), we get

$$Z(y-z, f(.)g(z|.)) = W(y-z, f(.)g(z|.)) - \Big(h(y-z)^+ + p(y-z)^- - c(y-z)\Big)\int f(\eta)g(z|\eta)d\eta. \tag{EC.43}$$

So,

$$\begin{aligned}
Z(x, f(.)) = \min_{y\geq x}\Bigg\{ & cy\int f(\eta)d\eta + \alpha\int_0^{+\infty} Z(y-z, f(.)g(z|.))dz \\
& + \alpha\int_0^{+\infty}\Big(h(y-z)^+ + p(y-z)^- - c(y-z)\Big)\int f(\eta)g(z|\eta)d\eta dz\Bigg\} \\
= \min_{y\geq x}\Bigg\{ & cy\int f(\eta)d\eta + \alpha\int_0^{+\infty} Z(y-z, f(.)g(z|.))dz \\
& + \alpha\int_0^{+\infty}\Big(h(y-z)^+ + p(y-z)^- - c(y-z)^+ + c(y-z)^-\Big)\int f(\eta)g(z|\eta)dz d\eta\Bigg\}
\end{aligned}$$

$$
\begin{aligned}
&= \min_{y \geq x} \Bigg\{ cy \int f(\eta)d\eta + \alpha \int_0^{+\infty} Z(y-z, f(.)g(z|.))dz \quad + \quad \underbrace{\underbrace{\alpha h \int_0^y (y-z) \int f(\eta) g(z|\eta) dz d\eta}_{= \alpha h \int_0^y (y-z) \int f(\eta) dG(z|\eta) d\eta}}_{=\alpha h \int_0^y \int f(\eta) G(z|\eta) dz d\eta = \alpha h y \int f(\eta) d\eta - \alpha h \int_0^y f(\eta) \int \bar{G}(z|\eta) dz d\eta} \\
&+ \underbrace{\underbrace{\alpha p \int_y^{+\infty} (y-z) \int f(\eta) g(z|\eta) dz d\eta}_{= -\alpha p \int_y^{+\infty} (y-z) \int f(\eta) d\bar{G}(z|\eta) d\eta}}_{=\alpha p \int_y^{+\infty} f(\eta) \bar{G}(z|\eta) d\eta} \underbrace{- \alpha cy \int_0^{+\infty} \int f(\eta) g(z|\eta) dz d\eta}_{=-\alpha cy \int f(\eta) d\eta} + \underbrace{\underbrace{\alpha c \int_0^{+\infty} z \int f(\eta) g(z|\eta) dz d\eta}_{= -\alpha c \int f(\eta) \int_0^{+\infty} z d\bar{G}(z|\eta) d\eta}}_{\alpha c \int f(\eta) \int_0^{+\infty} \bar{G}(z|\eta) dz d\eta = \alpha c \int f(\eta) \varphi(\eta) d\eta} \Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
&= \alpha c \int f(\eta) \varphi(\eta) d\eta + \min_{y \geq x} \Bigg\{ c(1-\alpha) y \int f(\eta) d\eta + \alpha \int_0^{+\infty} Z(y-z, f(.)g(z|.))dz + \alpha h y \int f(\eta) d\eta \\
&\qquad\qquad - \alpha h \int_0^y f(\eta) \int \bar{G}(z|\eta) d\eta + \alpha p \int_y^{+\infty} f(\eta) \bar{G}(z|\eta) d\eta \Bigg\} \\
&= \alpha(c-h) \int f(\eta) \varphi(\eta) d\eta + \min_{y \geq x} \Bigg\{ (c(1-\alpha) + \alpha h) y \int f(\eta) d\eta + \alpha \int_0^{+\infty} Z(y-z, f(.)g(z|.))dz \\
&\qquad\qquad + \alpha(p+h) \int_y^{+\infty} f(\eta) \bar{G}(z|\eta) d\eta \Bigg\}.
\end{aligned}
$$

□

EC.6. Proof of Proposition 2

Set

$$
\begin{aligned}
Q(Z)(y, f(.)) = (c(1-\alpha) + \alpha h) y \int f(\eta) d\eta + \alpha(h+p) \int_y^{+\infty} \int \bar{G}(z|\eta) f(\eta) d\eta dz \\
+ \alpha \int_0^{+\infty} Z(y-z, f(.)g(z|.))dz.
\end{aligned}
\tag{EC.44}
$$

The function $y \to Q(Z)(y, f(.))$ is differentiable in y, and denoting the derivative in y by $Q(Z)'(y, f(.))$, we have

$$
\begin{aligned}
Q(Z)'(y, f(.)) = (c(1-\alpha) + \alpha h) \int f(\eta) d\eta - \alpha(h+p) \int \bar{G}(y|\eta) f(\eta) d\eta \\
+ \alpha \int_0^{+\infty} Z'(y-z, f(.)g(z|.))dz.
\end{aligned}
$$

$$
Q(Z)'(y, f(.)) < 0, \text{ if } y \leq 0 \quad \text{and} \quad Q(Z)'(+\infty, f(.)) \geq (c(1-\alpha) + \alpha h) \int f(\eta) d\eta. \tag{EC.45}
$$

Therefore, there exists a single $S_Z(f(.))$ such that

$$Q(Z)'(S_Z(f(.)), f(.)) = 0. \tag{EC.46}$$

Necessarily $S_Z(f(.)) > 0$. Therefore, clearly

$$\Theta(Z)(x, f(.)) = \alpha(c-h)\int \varphi(\eta)f(\eta)d\eta + \left| \begin{array}{l} Q(Z)(S_Z(f(.)), f(.)), \quad \text{if } x < S_Z(f(.)), \\ \\ Q(Z)(x, f(.)), \quad \text{if } x \geq S_Z(f(.)), \end{array} \right. \tag{EC.47}$$

and

$$\Theta(Z)'(x, f(.)) = \left| \begin{array}{l} 0, \quad \text{if } x < S_Z(f(.)), \\ \\ Q(Z)'(x, f(.)), \quad \text{if } x \geq S_Z(f(.)). \end{array} \right.$$

We see that $\Theta(Z)'(x, f(.)) \geq 0$, and increasing in x. This proves that $\Theta(Z)(x, f(.))$ is increasing and convex. Moreover,

$$\sup_{x, f(.)} \frac{|Q(Z)'(x, f(.))|}{\int f(\eta)d\eta} \leq c(1-\alpha) + \alpha h + \alpha \sup_{x, f(.)} \frac{|Z'(x, f(.))|}{\int f(\eta)d\eta} < +\infty,$$

hence,

$$\sup_{x, f(.)} \frac{|\Theta(Z)'(x, f(.))|}{\int f(\eta)d\eta} < +\infty.$$

Also, $\Theta(Z)'(x, f(.)) = 0$ for $x \leq 0$. Hence (53) is also satisfied. The proof is completed. $\square$

EC.7. Proof of Theorem 2

We consider the increasing sequence

$$\left| \begin{array}{l} Z_{k+1}(x, f(.)) = \Theta(Z_k)(x, f(.)), \\ \\ Z_0(x, f(.)) = (h\dfrac{\alpha}{1-\alpha} + c)x^+ \displaystyle\int f(\eta)d\eta + \dfrac{\alpha(p+c)}{1-\alpha} \displaystyle\int \varphi(\eta)f(\eta)d\eta. \end{array} \right. \tag{EC.48}$$

Clearly $Z_0(x,f(.))$ is globally continuous and $\mathcal{C}^1$, increasing, and convex in x. It satisfies (52) and (53). By the stability properties of Proposition 2, we get sequentially the same properties for all the terms $Z_k(x,f(.))$, and therefore for the limit $Z(x,f(.))$. Also, we can check sequentially that

$$0 \leq \frac{(Z_k)'(x,f(.))}{\int f(\eta)d\eta} \leq \frac{c-\alpha c+\alpha h}{1-\alpha},$$

which carries to the limit

$$0 \leq \frac{Z'(x,f(.))}{\int f(\eta)d\eta} \leq \frac{c-\alpha c+\alpha h}{1-\alpha}. \tag{EC.49}$$

The optimal feedback is then defined by a base stock policy, with a base stock $S_Z(f.))$ unique solution of (EC.46). This complete the proof. □

EC.7.1. Proof of Proposition 3

Since $(c(1-\alpha)+\alpha h)\displaystyle\int f(\eta)d\eta - \alpha(h+p)\int \bar{G}(x|\eta)f(\eta)d\eta$ satisfies (58), we see easily that the contraction mapping theorem can be applied for the fixed point equation (55), hence the result.

□

EC.8. Computation of $Z_2'(x,\beta)$

We have

$$\begin{aligned} Z_2'(x,\beta) = \Big(& (c(1-\alpha)+\alpha h)(\beta+1) - \alpha(h+p)(\beta e^{-\delta_0 x^+} + e^{-\delta_1 x^+}) \\ & + \alpha\int_0^{+\infty} \delta_1 e^{-\delta_1 z} Z_1'\Big(x-z, \beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big)dz\Big)^+, \end{aligned} \tag{EC.50}$$

so we need to compute first

$$Z_1'\big(x-z, \beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\big), \ \ \forall z>0. \tag{EC.51}$$

From (101) we have

$$Z_1'\big(x-z,\beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\big) = \left| \begin{array}{l} 0, \text{ if } x-z < S_1\big(\beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\big), \\ \\ (c(1-\alpha)+\alpha h)\big(\beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}+1\big) \\ -\alpha(h+p)\big(\beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}e^{-\delta_0(x-z)}+e^{-\delta_1(x-z)}\big), \text{ if } x-z \geq S_1\big(\beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\big). \end{array} \right. \tag{EC.52}$$

We can define

$$s_1(z) = z + S_1\Big(\beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big). \tag{EC.53}$$

The function $s_1(z)$ is monotone increasing and $s_1(0) = S_1\Big(\beta\frac{\delta_0}{\delta_1}\Big)$, $s_1(\infty) = \infty$. We can see that if $x < S_1\Big(\beta\frac{\delta_0}{\delta_1}\Big) < S_1(\beta)$, $Z_2'(x,\beta) = 0$. We may assume that $x > S_1\Big(\beta\frac{\delta_0}{\delta_1}\Big)$. There exists a single value $z_1(x)$ such that

$$s_1(z_1(x)) = x, \quad \text{and} \quad x > s_1(z) \Longleftrightarrow z < z_1(x). \tag{EC.54}$$

Note that $z_1\Big(S_1(\beta\frac{\delta_0}{\delta_1})\Big) = 0$, so we may set $z_1(x) = 0$, if $x < S_1\big(\beta\frac{\delta_0}{\delta_1}\big)$. The function $z_1(x)$ is monotone increasing. From (EC.52) we can write

$$Z_1'\Big(x-z, \beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big) = \left|\begin{array}{l} 0, \ \text{if} \ z > z_1(x), \\ \\ (c(1-\alpha)+\alpha h)\Big(\beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}+1\Big) - \alpha(h+p)e^{\delta_1 z}\Big(\beta\frac{\delta_0}{\delta_1}e^{-\delta_0 x}+e^{-\delta_1 x}\Big), \\ \ \text{if} \ z \le z_1(x). \end{array}\right.$$

Therefore,

$$\begin{aligned} &\alpha\int_0^{+\infty}\delta_1 e^{-\delta_1 z}Z_1'\Big(x-z,\beta\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big)dz \\ &= \alpha\int_0^{z_1(x)}\Big[(c(1-\alpha)+\alpha h)\big(\beta\delta_0 e^{-\delta_0 z}+\delta_1 e^{-\delta_1 z}\big) - \alpha(h+p)\big(\beta\delta_0 e^{-\delta_0 x}+\delta_1 e^{-\delta_1 x}\big)\Big]dz \\ &= \alpha\Big[(c(1-\alpha)+\alpha h)\Big(\beta\big(1-e^{-\delta_0 z_1(x)}\big)+1-e^{-\delta_1 z_1(x)}\Big) - \alpha(h+p)\Big(\beta\delta_0 e^{-\delta_0 x}+\delta_1 e^{-\delta_1 x}\Big)z_1(x)\Big]. \end{aligned} \tag{EC.55}$$

So we get

$$\begin{aligned} Z_2'(x,\beta) = \Big(&(c(1-\alpha)+\alpha h)\big[\beta\big(1+\alpha(1-e^{-\delta_0 z_1(x)})\big)+1+\alpha(1-e^{-\delta_1 z_1(x)})\big] \\ &-\alpha(h+p)\Big[e^{-\delta_0 x^+}\beta\big(1+\alpha\delta_0 z_1(x)\big)+e^{-\delta_1 x^+}\big(1+\alpha\delta_1 z_1(x)\big)\Big]\Big)^+. \end{aligned} \tag{EC.56}$$

The function

$$\begin{aligned} &(c(1-\alpha)+\alpha h)\big[\beta\big(1+\alpha(1-e^{-\delta_0 z_1(x)})\big)+1+\alpha(1-e^{-\delta_1 z_1(x)})\big] \\ &-\alpha(h+p)\Big[e^{-\delta_0 x}\beta\big(1+\alpha\delta_0 z_1(x)\big)+e^{-\delta_1 x}\big(1+\alpha\delta_1 z_1(x)\big)\Big] \end{aligned} \tag{EC.57}$$

is continuous monotone increasing from $(c(1-\alpha)+\alpha h)(\beta+1)-\alpha(h+p)(\beta+1)<0$, for $x=-\infty$ to $(c(1-\alpha)+\alpha h)(\beta+1)(1+\alpha)$, for $x=+\infty$. Hence, there exists a unique $x=S_2(\beta)>0$ such that

$$\begin{aligned}&(c(1-\alpha)+\alpha h)\left[\beta\left(1+\alpha(1-e^{-\delta_0 z_1(x,\beta)})\right)+1+\alpha(1-e^{-\delta_1 z_1(x,\beta)})\right]\\&-\alpha(h+p)\Big[e^{-\delta_0 x}\beta\left(1+\alpha\delta_0 z_1(x,\beta)\right)+e^{-\delta_1 x}\left(1+\alpha\delta_1 z_1(x,\beta)\right)\Big]=0,\end{aligned} \tag{EC.58}$$

where we have reinserted the fact that $z_1(x)$ depends on β. Finally,

$$Z_2'(x,\beta)=\left|\begin{array}{l}0, \quad \text{if} \quad x<S_2(\beta),\\ \\ (c(1-\alpha)+\alpha h)\left[\beta\left(1+\alpha(1-e^{-\delta_0 z_1(x,\beta)})\right)+1+\alpha(1-e^{-\delta_1 z_1(x,\beta)})\right]\\ -\alpha(h+p)\Big[e^{-\delta_0 x}\beta\left(1+\alpha\delta_0 z_1(x,\beta)\right)+e^{-\delta_1 x}\left(1+\alpha\delta_1 z_1(x,\beta)\right)\Big], \quad \text{if} \quad x\geq S_2(\beta).\end{array}\right. \tag{EC.59}$$

Consider next the case where $\dfrac{\delta_0}{\delta_1}=2$ and $\delta_1=\gamma$. Then we can state, see (100)

$$e^{-\gamma S_1(\beta)}=\frac{-1+\sqrt{1+4\beta(\beta+1)\frac{c(1-\alpha)+\alpha h}{\alpha(h+p)}}}{2\beta}. \tag{EC.60}$$

Hence,

$$S_1(\beta)=\frac{1}{\gamma}\log\frac{2\beta}{-1+\sqrt{1+4\beta(\beta+1)\frac{c(1-\alpha)+\alpha h}{\alpha(h+p)}}}. \tag{EC.61}$$

And thus, from (EC.53),

$$s_1(z)=z+\frac{1}{\delta}\log\frac{4\beta e^{-\gamma z}}{-1+\sqrt{1+8\beta e^{-\gamma z}(1+2\beta e^{-\gamma z})\frac{c(1-\alpha)+\alpha h}{\alpha(h+p)}}}. \tag{EC.62}$$

For $x>S_1(2\beta)$, we get the formula

$$z_1(x)=\frac{1}{\gamma}\log\frac{4\beta}{-1+\sqrt{1+8\beta e^{-\gamma x}(1+2\beta e^{-\gamma x})\frac{\alpha(h+p)}{c(1-\alpha)+\alpha h}}}. \tag{EC.63}$$

Since $Z_1'(x,2\beta)>0$ for $x>S_1(2\beta)$, the condition for x means

$$2\beta e^{-2\gamma x}+e^{-\gamma x}<(1+2\beta)\frac{c(1-\alpha)+\alpha h}{\alpha(h+p)}, \tag{EC.64}$$

and equation (EC.58) reduces to

$$\begin{aligned}&(c(1-\alpha)+\alpha h)\Big[(1+\alpha)(1+\beta)-\alpha\beta e^{-2\gamma z_1(x,\beta)}-\alpha e^{-\gamma z_1(x,\beta)}\Big]\\&-\alpha(h+p)\Big[\beta(1+2\alpha\gamma z_1(x,\beta))e^{-2\gamma x}+(1+\alpha\gamma z_1(x,\beta))e^{-\gamma x}\Big]=0.\end{aligned}\tag{EC.65}$$

It is convenient to solve (EC.65) in terms of x as a function of z. We obtain

$$x(z_1(x))=\frac{1}{\gamma}\log\frac{2\beta(1+2\alpha\gamma z_1(x))}{-(1+\alpha\gamma z_1(x))+\sqrt{f(z_1(x))}},\tag{EC.66}$$

where $f(z_1(x))$ is given by

$$f(z_1(x))=(1+\alpha\gamma z_1(x))^2+\frac{4\beta(1+2\alpha\gamma z_1(x))(c(1-\alpha)+\alpha h)\big[(1+\alpha)(1+\beta)-\alpha\beta e^{-2\gamma z_1(x)}-\alpha e^{-\gamma z_1(x)}\big]}{\alpha(h+p)},\tag{EC.67}$$

and $S_2(\beta)$ solves the fixed point equation $x=x(z_1(x))$. □

EC.9. Computation of $Z_0'(x)$, $Z_1'(x)$, $Z_2'(x)$ and $Z_3'(x)$

We approximate the function $Z'(x,\beta)$ by a piecewise constant function in the parameter β, so we write

$$Z'(x,\beta)\approx Z'_{[\beta]}(x),\ \ [\beta]\le\beta<[\beta]+1,\tag{EC.68}$$

where $[\beta]$ is the integer part of β. We can therefore approximate (97) as

$$\begin{aligned}Z_k'(x)=\Big(&(c(1-\alpha)+\alpha h)(k+1)-\alpha(h+p)(ke^{-\delta_0 x^+}+e^{-\delta_1 x^+})\\&+\alpha\delta_1\int_0^{+\infty}e^{-\delta_1 z}Z'\Big(x-z,k\frac{\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big)dz\Big)^+,\end{aligned}\tag{EC.69}$$

where k is an integer, and find it by by approximating the integral on its right hand side. We have

$$\left|\begin{aligned}&0\le z<\frac{1}{\delta_0-\delta_1}\log\frac{\frac{k\delta_0}{\delta_1}}{\big[\frac{k\delta_0}{\delta_1}\big]}\Longrightarrow Z'\Big(y,\frac{k\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big)\approx Z'_{\big[\frac{k\delta_0}{\delta_1}\big]}(y),\\&\frac{1}{\delta_0-\delta_1}\log\frac{\frac{k\delta_0}{\delta_1}}{\big[\frac{k\delta_0}{\delta_1}\big]-j}\le z<\frac{1}{\delta_0-\delta_1}\log\frac{\frac{k\delta_0}{\delta_1}}{\big[\frac{k\delta_0}{\delta_1}\big]-(j+1)}\Longrightarrow Z'\Big(y,\frac{k\delta_0}{\delta_1}e^{-(\delta_0-\delta_1)z}\Big)\approx Z'_{\big[\frac{k\delta_0}{\delta_1}\big]-(j+1)}(y),\\&0\le j\le\Big[\frac{k\delta_0}{\delta_1}\Big]-1,\end{aligned}\right.\tag{EC.70}$$

and with that (EC.69) becomes

$$
\begin{aligned}
Z_k'(x) = \Bigg(& (c(1-\alpha)+\alpha h)(k+1) - \alpha(h+p)(ke^{-\delta_0 x^+} + e^{-\delta_1 x^+}) \\
& + \alpha\delta_1 \int_0^{\frac{1}{\delta_0-\delta_1}(\log\frac{k\delta_0}{\delta_1} - \log[\frac{k\delta_0}{\delta_1}])} e^{-\delta_1 z} Z'_{\left[\frac{k\delta_0}{\delta_1}\right]}(x-z)dz \\
& + \alpha\delta_1 \sum_{j=0}^{[\frac{k\delta_0}{\delta_1}]-1} \int_{\frac{1}{\delta_0-\delta_1}(\log\frac{k\delta_0}{\delta_1} - \log([\frac{k\delta_0}{\delta_1}]-j))}^{\frac{1}{\delta_0-\delta_1}(\log\frac{k\delta_0}{\delta_1} - \log([\frac{k\delta_0}{\delta_1}]-j-1))} e^{-\delta_1 z} Z'_{\left[\frac{k\delta_0}{\delta_1}\right]-j-1}(x-z)dz \Bigg)^+.
\end{aligned}
\tag{EC.71}
$$

We will focus on the approximations of $Z_0'(x)$, $Z_1'(x)$, and $Z_2'(x)$, and see how the algorithm evolves. We cannot use (EC.71) with $k=0$, which corresponds to the case that the true distribution is $g_1(x)$. In this case there is no belief, and (97) yields

$$
Z_0'(x) = \Big(c(1-\alpha) + \alpha h - \alpha(h+p)e^{-\delta_1 x^+} + \alpha \int_0^{+\infty} \delta_1 e^{-\delta_1 z} Z_0'(x-z)dz \Big)^+. \tag{EC.72}
$$

The solution of (EC.72) is explicit (see Section 3), namely,

$$
Z_0'(x) = \left| \begin{array}{l}
0, \text{ if } x \le S_0, \text{ where } e^{-\delta_1 S_0} = \dfrac{c(1-\alpha)+\alpha h}{\alpha(h+p)}, \\[2em]
\dfrac{1}{1-\alpha}\Big(c(1-\alpha)+\alpha h - \alpha(h+p)e^{-(1-\alpha)\delta_1 x} e^{-\alpha\delta_1 S_0} \Big), \text{ if } x \ge S_0.
\end{array} \right. \tag{EC.73}
$$

From (EC.71) we see that $Z_k'(x)$ depends on $Z_0'(x), \cdots, Z'_{\left[\frac{k\delta_0}{\delta_1}\right]-1}(x)$. There is a slight simplification when δ_0/δ_1 is an integer, in which case $Z_k'(x)$ depends on $Z_0'(x), \cdots, Z'_{\left[\frac{k\delta_0}{\delta_1}\right]-1}(x)$. For instance, suppose $\delta_0/\delta_1 = 2$, in which case $Z_k'(x)$ depends on $Z_0'(x), \cdots, Z_{2k-1}'(x)$. Then by setting $\delta_1 = \gamma$,

$$
\begin{aligned}
Z_k'(x) = \Big(& (c(1-\alpha)+\alpha h)(k+1) - \alpha(h+p)(ke^{-2\gamma x^+} + e^{-\gamma x^+}) \\
& + \alpha\gamma \sum_{j=0}^{2k-1} \int_{\frac{1}{\gamma}(\log 2k - \log(2k-j))}^{\frac{1}{\gamma}(\log 2k - \log(2k-j-1))} e^{-\gamma z} Z'_{2k-j-1}(x-z)dz \Big)^+.
\end{aligned}
\tag{EC.74}
$$

We can in particular compute $Z_1'(x)$ by solving the equation

$$
\begin{aligned}
Z_1'(x) = \Big(& 2(c(1-\alpha)+\alpha h) - \alpha(h+p)(e^{-2\gamma x^+} + e^{-\gamma x^+}) + \alpha\gamma \int_0^{\frac{\log 2}{\gamma}} e^{-\gamma z} Z_1'(x-z)dz \\
& + \alpha\gamma \int_{\frac{\log 2}{\gamma}}^{+\infty} e^{-\gamma z} Z_0'(x-z)dz \Big)^+.
\end{aligned}
\tag{EC.75}
$$

From (EC.73) we compute $\alpha\gamma\int_{\frac{\log 2}{\gamma}}^{+\infty} e^{-\gamma z} Z_0'(x-z)dz$. It is 0 if $x \le S_0 + \frac{\log 2}{\gamma}$. For $x > S_0 + \frac{\log 2}{\gamma}$, we obtain, after easy calculations,

$$\begin{aligned}\alpha\gamma\int_{\frac{\log 2}{\gamma}}^{+\infty} e^{-\gamma z} Z_0'(x-z)dz &= \alpha\gamma\int_{\frac{\log 2}{\gamma}}^{x-S_0} e^{-\gamma z} Z_0'(x-z)dz \\ &= \frac{\alpha}{1-\alpha} e^{-\gamma x}\Big((c(1-\alpha)+\alpha h)\big(\frac{e^{\gamma x}}{2} - e^{\gamma S_0}\big) - (h+p)\big(\frac{e^{\gamma\alpha(x-S_0)}}{2^\alpha} - 1\big)\Big).\end{aligned}$$

By the definition of S_0 and (EC.73), we obtain

$$\alpha\gamma\int_{\frac{\log 2}{\gamma}}^{x-S_0} e^{-\gamma z} Z_0'(x-z)dz = \frac{\alpha}{1-\alpha}(h+p)e^{-\gamma x}\Big(1-\alpha+\frac{\alpha e^{\gamma(x-S_0)}}{2} - \frac{e^{\alpha\gamma(x-S_0)}}{2^\alpha}\Big), \ \forall x > S_0 + \frac{\log 2}{\gamma}, \tag{EC.76}$$

and we check immediately that the expression (EC.76) is positive. We next define S_1 such that

$$2(c(1-\alpha)+\alpha h) - \alpha(h+p)(e^{-2\gamma S_1} + e^{-\gamma S_1}) = 0, \tag{EC.77}$$

and we clearly have $S_1 < S_0$. Then we can write equivalently for (EC.75) that

$$Z_1'(x) = \left|\begin{array}{l} 0, \ \text{if } \ x < S_1, \\[2em] 2(c(1-\alpha)+\alpha h) - \alpha(h+p)(e^{-2\gamma x} + e^{-\gamma x}) + \alpha\gamma\displaystyle\int_0^{\frac{\log 2}{\gamma}} e^{-\gamma z} Z_1'(x-z)dz, \\ \text{if } \ S_1 \le x < S_0 + \dfrac{\log 2}{\gamma}, \\[2em] 2(c(1-\alpha)+\alpha h) - \alpha(h+p)e^{-2\gamma x} + \dfrac{\alpha}{1-\alpha}(h+p)e^{-\gamma x}\Big(\dfrac{\alpha e^{\gamma(x-S_0)}}{2} - \dfrac{e^{\alpha\gamma(x-S_0)}}{2^\alpha}\Big) \\ +\alpha\gamma\displaystyle\int_0^{\frac{\log 2}{\gamma}} e^{-\gamma z} Z_1'(x-z)dz, \ \text{if } \ x \ge S_0 + \dfrac{\log 2}{\gamma}. \end{array}\right. \tag{EC.78}$$

Indeed, the integral equation (EC.78) has a unique positive, continuous, bounded, and monotone increasing solution $Z_1'(x)$. Unfortunately, unlike for $Z_0'(x)$, we cannot give a closed-form formula for $Z_1'(x)$. On the other hand, using a composite trapezoidal formula to approximate the integral in (EC.78), i.e.,

$$\int_0^{\frac{\log 2}{\gamma}} e^{-\gamma z} Z_1'(x-z)dz \approx \frac{\log 2}{2\gamma N}\Big[Z_1'(x) + 2\sum_{k=1}^{N-1} e^{-\gamma z_k} Z_1'(x-z_k) + Z_1'(x - \frac{\log 2}{\gamma})\Big], \tag{EC.79}$$

where $z_k = \dfrac{\log 2}{\gamma N} k,\ k = 1, \cdots, N-1$, we get the algebraic equation

$$Z_1'(x) = W(x) + \alpha \frac{\log 2}{2N}\Big[Z_1'(x) + 2\sum_{k=1}^{N-1} e^{-\gamma z_k} Z_1'(x - z_k) + Z_1'(x - \frac{\log 2}{\gamma})\Big], \tag{EC.80}$$

where

$$W(x) = \left| \begin{array}{l} 2(c(1-\alpha)+\alpha h) - \alpha(h+p)(e^{-2\gamma x} + e^{-\gamma x}),\ \text{ if }\ S_1 \le x < S_0 + \dfrac{\log 2}{\gamma}, \\ \\ 2(c(1-\alpha)+\alpha h) - \alpha(h+p)e^{-2\gamma x} + \dfrac{\alpha}{1-\alpha}(h+p)e^{-\gamma x}\Big(\dfrac{\alpha e^{\gamma(x-S_0)}}{2} - \dfrac{e^{\alpha\gamma(x-S_0)}}{2^\alpha}\Big), \\ \text{if }\ x \ge S_0 + \dfrac{\log 2}{\gamma}. \end{array} \right.$$

Using (EC.80) as a forward marching scheme, we can obtain the value of $Z_1'(x)$ at the points $S_1 + \dfrac{\log 2}{\gamma N} k,\ k = 1, 2, \cdots$. We define $Z_1'(x)$ for all finite x as the piecewise linear function that is 0 for $x < S_1$, and connecting the above data points for $x \ge S_1$. Note that (EC.80) yields the limiting relation $Z_1'(\infty) = \dfrac{W(\infty)}{1-\alpha/2}$.

Unfortunately, the computation of $Z_2'(x)$ cannot be obtained by a single equation. Indeed, from (EC.74) we have

$$\begin{aligned} Z_2'(x) = \Big(&3(c(1-\alpha)+\alpha h) - \alpha(h+p)(2e^{-2\gamma x^+} + e^{-\gamma x^+}) + \alpha\gamma\Big[\int_{\frac{\log 4}{\gamma}}^{+\infty} e^{-\gamma z} Z_0'(x-z)dz \\ &+ \int_{\frac{\log 2}{\gamma}}^{\frac{\log 4}{\gamma}} e^{-\gamma z} Z_1'(x-z)dz + \int_{\frac{\log \frac{4}{3}}{\gamma}}^{\frac{\log 2}{\gamma}} e^{-\gamma z} Z_2'(x-z)dz + \int_0^{\frac{\log \frac{4}{3}}{\gamma}} e^{-\gamma z} Z_3'(x-z)dz\Big]\Big)^+. \end{aligned} \tag{EC.81}$$

The right hand side of (EC.81) depends on $Z_3'(x)$, which cannot be computed before $Z_2'(x)$. If we write the equation for $Z_3'(x)$, it will involve $Z_4'(x)$ and $Z_5'(x)$ which are not known. To close the chain we take $Z_4'(x)$ and $Z_5'(x)$ equal to $Z_3'(x)$. So we write the equation $Z_3'(x)$ as follows:

$$\begin{aligned} Z_3'(x) = \Big(&4(c(1-\alpha)+\alpha h) - \alpha(h+p)(3e^{-2\gamma x^+} + e^{-\gamma x^+}) + \alpha\gamma\Big[\int_{\frac{\log 6}{\gamma}}^{+\infty} e^{-\gamma z} Z_0'(x-z)dz \\ &+ \int_{\frac{\log 3}{\gamma}}^{\frac{\log 6}{\gamma}} e^{-\gamma z} Z_1'(x-z)dz + \int_{\frac{\log 2}{\gamma}}^{\frac{\log 3}{\gamma}} e^{-\gamma z} Z_2'(x-z)dz + \int_0^{\frac{\log 2}{\gamma}} e^{-\gamma z} Z_3'(x-z)dz\Big]\Big)^+. \end{aligned} \tag{EC.82}$$

We obtain in this way a system for $Z_2'(x)$ and $Z_3'(x)$. It can be slightly simplified, considering S_3 to be the solution of

$$4(c(1-\alpha)+\alpha h)-\alpha(h+p)(3e^{-2\gamma S_3}+e^{-\gamma S_3})=0. \tag{EC.83}$$

We then have to find S_2, Z_2', Z_3' satisfying $S_3<S_2<S_1$ and

$$3(c(1-\alpha)+\alpha h)-\alpha(h+p)(2e^{-2\gamma S_2}+e^{-\gamma S_2})+\alpha\gamma\int_0^{\frac{\log\frac{4}{3}}{\gamma}} e^{-\gamma z}Z_3'(S_2-z)dz=0, \tag{EC.84}$$

$$Z_2'(x)=\left|\begin{array}{l} 0,\ \text{ if } \ x<S_2,\\ \\ 3(c(1-\alpha)+\alpha h)-\alpha(h+p)(2e^{-2\gamma x}+e^{-\gamma x})+\alpha\gamma\Big(\displaystyle\int_{\frac{\log 4}{\gamma}}^{+\infty} e^{-\gamma z}Z_0'(x-z)dz\\ +\displaystyle\int_{\frac{\log 2}{\gamma}}^{\frac{\log 4}{\gamma}} e^{-\gamma z}Z_1'(x-z)dz+\int_{\frac{\log\frac{4}{3}}{\gamma}}^{\frac{\log 2}{\gamma}} e^{-\gamma z}Z_2'(x-z)dz+\int_0^{\frac{\log\frac{4}{3}}{\gamma}} e^{-\gamma z}Z_3'(x-z)dz\Big),\ \ x\ge S_2,\end{array}\right. \tag{EC.85}$$

$$Z_3'(x)=\left|\begin{array}{l} 0,\ \text{ if } \ x<S_3,\\ \\ 4(c(1-\alpha)+\alpha h)-\alpha(h+p)(3e^{-2\gamma x}+e^{-\gamma x})+\alpha\gamma\Big(\displaystyle\int_{\frac{\log 6}{\gamma}}^{+\infty} e^{-\gamma z}Z_0'(x-z)dz\\ +\displaystyle\int_{\frac{\log 3}{\gamma}}^{\frac{\log 6}{\gamma}} e^{-\gamma z}Z_1'(x-z)dz+\int_{\frac{\log 2}{\gamma}}^{\frac{\log 3}{\gamma}} e^{-\gamma z}Z_2'(x-z)dz+\int_0^{\frac{\log 2}{\gamma}} e^{-\gamma z}Z_3'(x-z)dz\Big),\ \ x\ge S_3.\end{array}\right. \tag{EC.86}$$

Remark EC.1. As a very crude approximation to obtain S_2, we may approximate in equation (EC.86) the function $Z_3'(x)$ by

$$Z_3'(x)=\left|\begin{array}{l} 0,\ \text{ if } \ x<S_3,\\ \\ 4(c(1-\alpha)+\alpha h)-\alpha(h+p)(3e^{-2\gamma x}+e^{-\gamma x})\\ +\alpha\gamma\Big(\displaystyle\int_{\frac{\log 6}{\gamma}}^{+\infty} e^{-\gamma z}Z_0'(x-z)dz+\int_{\frac{\log 3}{\gamma}}^{\frac{\log 6}{\gamma}} e^{-\gamma z}Z_1'(x-z)dz\Big),\ \ x\ge S_3.\end{array}\right. \tag{EC.87}$$

Of course, we loose the continuity at point S_3. We can compute S_3 from (EC.83) and then $Z_3'(x)$ from (EC.87) using $Z_0'(x)$ and $Z_1'(x)$, (and a process similar to (EC.79) to approximate the second integral in (EC.87)). Knowing $Z_3'(x)$, we go back to (EC.84) and (EC.85) and find respectively the approximation of S_2 and $Z_2'(x)$.

REMARK EC.2. It is interesting to point out that the above approximation should not be considered as an approximation for the pair $\{Z_2'(x),\ Z_3'(x)\}$, but for $Z_2'(x)$ and S_2 only. We have approximated $Z_3'(x)$ and S_3 only in the context of $Z_2'(x)$ and S_2.

□

EC.10. Some Observations of the Approximations (9.1.1) and (9.1.3)

The approximations defined in Subsections 9.1.1 and 9.1.3 produce functions of one variable and two variables, respectively. Based on the numerical experiments, it is our objective to compare the two computational approaches with $\beta = 0, 1, 2$. First, if $\beta = 0$, there is no learning and (EC.72) yields exactly (98). Second, if $\beta = 1, 2$, we graph $Z_\beta'(x)$ and $Z'(x, \beta)$ through the data points $x_n = n\Delta x$ with $N = 600$, and see how the values of the functions compare.

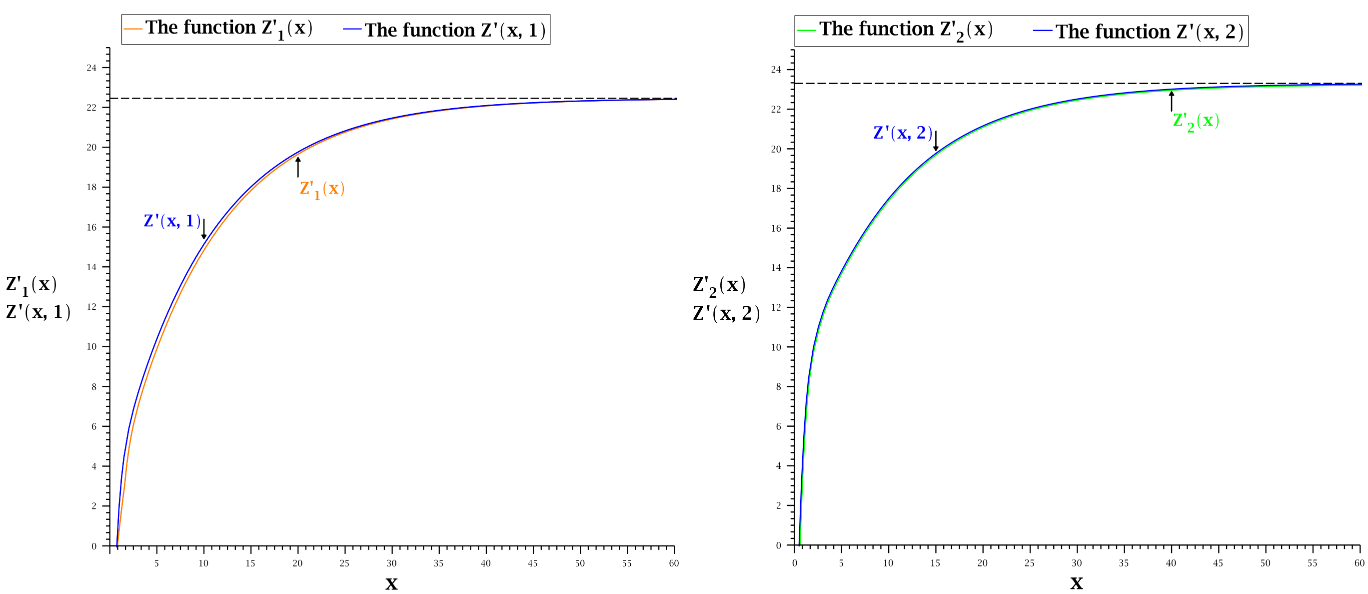

Figure EC.1 **The functions $Z_\beta'(x)$ and $Z'(x, \beta)$, $\beta = 1, 2$.**

We can see that $Z'_\beta(x)$ and $Z'(x,\beta)$ are very close for $\beta = 1,2$. In the implementation of the approximation by iteration, we first use the analytical expression (EC.59) of $Z'_2(x,\beta)$ to see how a change to a finer grid improves the accuracy of our results. Our numerical simulation shows that $\Delta x = 0.1$, $\Delta\beta = 0.05$ give us good grids and so they can be used in the grid point values for (105). Next, we resume the values S_β and $S(\beta)$ obtained for $\beta = 0,1,2,3$.

Table EC.1 **Approximation for** $\beta = 0,1,2,3$.

β	0	1	2	3
S_β	0.8622	0.6048	0.5370	0.5076
$S(\beta)$	0.8622	0.6012	0.5351	0.5058
$\lvert S_\beta - S(\beta)\rvert$	0	0.0036	0.0019	0.0018

Our numerical results indicate that $S_\beta \approx S(\beta)$. Note that $S(0) = S_0$ (no learning if $\beta = 0$). Also,

$$|S_3 - S(3)| < |S_2 - S(2)| < |S_1 - S(1)|. \tag{EC.88}$$

For $k > 2$, we have to approximate the solution of the functional equation by computational means, e.g., by applying piecewise constant approximations or preferably, by using successive approximations.

Acknowledgments

Alain Bensoussan gratefully acknowledges the support of the National Science Foundation under grants DMS 1612880 and DMS 1905449.

References

Azoury KS (1985) Bayes solution to dynamic inventory models under unknown demand distribution. *Management. Sci.* **31** 1150-1160.

Bensoussan A, Crouhy M, Proth JM (1983) *Mathematical Theory of Production Planning.* (North-Holland, New York).

Bensoussan A, Çakanyildirim M, Sethi SP (2007) A multiperiod newsvendor problem with partially observed demand. *Math. Oper. Res.* **32**(2) 322-344.

Bensoussan A, Çakanyildirim M, Minjarez-Sosa JA, Royal A, Sethi SP (January 2008) Inventory problems with partially observed demands and lost sales. *J. Optim. Theory Appl.* **136**(3) 321-340.

Bensoussan A, Çakanyildirim M, Minjarez-Sosa JA, Sethi SP, Shi R (February 2008) Partially observed inventory systems: The case of rain checks. *SIAM J. Control Optim.* **47**(5) 2490-2519.

Bensoussan A, Çakanyildirim M, Sethi SP (2009a) A Note on "The censored newsvendor and the optimal acquisition of information". *Oper. Res.* **57**(3) 791-794.

Bensoussan A, Çakanyildirim M, Royal A, Sethi SP (2009c) Bayesian and adaptive controls for a newsvendor facing exponential demand. *J. Risk Decision Anal.* **1**(4) 197-210.

Bensoussan A, Çakanyildirim M, Sethi SP (2010) Filtering for discrete-time Markov processes and applications to inventory control with incomplete information. *Handbook on Nonlinear Filtering*, Crisan D and Rozovsky B, eds. (Oxford University Press), 500-525.

Bensoussan A (2011) *Dynamic Programming and Inventory Control,* Studies in Probability, Optimization and Statistics, Vol. 3. (IOS Press, Amsterdam).

Bensoussan A, Çakanyildirim M, Li M, Sethi SP (2016) Managing inventory with cash register information: Sales recorded but not demands. *Production Oper. Management* **25**(1) 9-21.

Bisi A, Dada M (2007) Dynamic learning, pricing, and ordering by a censored newsvendor. *Naval Logist. Res.* **54**(4) 448-461.

Brezis H (2011) *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, (Springer, Universitext, New York).

Braden DJ, Freimer M (1991) Informational dynamics of censored observations. *Management. Sci.* **37** 1390-1404.

Chen L, Plambeck EL (2008) Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing Service Oper. Management* **10**(2) 236-256.

Ding X, Puterman ML, Bisi A (2002) The censored newsvendor and the optimal acquisition of information. *Oper. Res.* **50**(3) 517-527.

Hakzos C, Seshadri S (2004) Monotone forecasts. *Oper. Res.* **52** 478-486.

Harpaz G, Lee W, Winkler R (1982) Learning, experimentation, and the optimal output decisions of a competitive firm. *Management Sci.* **28**(6) 589-603.

Lariviere MA, Porteus EL (1999) Stalking information: Bayesian inventory management with unobserved lost sales. *Management. Sci.* **45**(3) 346-363.

Lu X, Song JS, Zhu K (2005) On "The censored newsvendor and the optimal acquisition of information. *Oper. Res.* **53** 1024-1027.

Miller B (1986) Scarf's state reduction method, flexibility, and a dependent demand inventory model. *Oper. Res.* **34** 83-90.

Murray Jr. G, Silver E (1966) A bayesian analysis of the style goods inventory problem. *Management. Sci.* **12**(11) 785-797.

Royden HL (1968) *Real Analysis.* Second Edition, (The Macmillan Company, New York).

Scarf H (1960a) Some remarks on Bayes solutions to the inventory problem. *Naval Res. Logist. Quart.* **7** 591-596.

Sethi SP (October 2010) i3: Incomplete information inventory models. *Decision Line* **41**(5) 16-19.

Treharne JT, Sox CR (2002) Inventory control for nonstationary demand and partial observation. *Management. Sci.* **48** 607-624.