

## OPTIMAL STRUCTURAL POLICIES FOR AMBIGUITY AND RISK AVERSE INVENTORY AND PRICING MODELS\*

XIN CHEN<sup>†</sup> AND PENG SUN<sup>‡</sup>

**Abstract.** This paper discusses multiperiod stochastic joint inventory and pricing models when the decision maker is risk and ambiguity averse. We study infinite horizon models with discounted and long run average optimization criteria. The main result of this paper is establishing the optimality of stationary  $(s, S, p)$  policies for the infinite horizon inventory and pricing models.

**Key words.** risk averse, ambiguity averse, inventory and pricing, infinite horizon dynamic program,  $(s, S)$  policy

**AMS subject classifications.** 90B05, 90B60

**DOI.** 10.1137/100791488

**1. Introduction.** Risk neutrality and complete knowledge of demand probability distributions are two underlying assumptions behind most conventional stochastic dynamic inventory control models. In practice, however, these assumptions do not always hold and may be too restrictive. On the one hand, an inventory manager may be risk averse, i.e., she prefers a control policy that protects the downside risk at the expense of the average performance. On the other hand, a decision maker may not know the exact demand distributions and have to estimate them from limited historical data. In this case, the decision maker tends to be ambiguity averse, i.e., she prefers an inventory policy which is “robust” against estimation errors.

In this paper we consider a single product, periodic review inventory system with a *fixed ordering cost* and stochastic price-dependent demand. The objective is to make inventory and pricing decisions over time so as to maximize a certain performance measure that takes into account risk and ambiguity aversion. Our main contribution is to show that for the proposed ambiguity and risk averse inventory and pricing model, a stationary  $(s, S, p)$  policy is optimal in the infinite horizon setting, although a weaker structure,  $(s, S, A, p)$  policy, is optimal for its finite horizon counterpart.<sup>1</sup>

Establishing optimal structural policies for infinite horizon inventory models with fixed ordering cost is a challenging problem, which motivated many classic papers in the stochastic inventory control literature (see Iglehart (1963) and Zheng (1991) as well as the references therein). More recently, there is an active research stream on supply chain models with risk averse agents (see, for example, Chen et al. (2007), Gan, Sethi, and Yan (2004), and Sobel and Turcic (2008), as well as the references therein). Incorporating ambiguity aversion into operations models has also drawn attentions

---

\*Received by the editors April 8, 2010; accepted for publication (in revised form) September 27, 2011; published electronically January 5, 2012.

<http://www.siam.org/journals/sicon/50-1/79148.html>

<sup>†</sup>Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 (xinchen@illinois.edu). This author’s research was partly supported by the National Science Foundation grants CMMI-0653909, CMMI-0926845 ARRA, and CMMI-1030923.

<sup>‡</sup>Fuqua School of Business, Duke University, Box 90120, Durham, NC 27708 (psun@duke.edu).

<sup>1</sup>In an  $(s, S, A, p)$  policy, at each period there exist two parameters  $s$  and  $S$  as well as a set  $A \in (s, \frac{s+S}{2}]$ , such that an order is placed if the initial inventory level is no more than  $s$  or belongs to  $A$ ; otherwise no order is placed. The optimal price is a function of the initial inventory level. An  $(s, S, p)$  policy is an  $(s, S, A, p)$  with  $A$  being empty.

from the stochastic optimal control community (see Lim and Shanthikumar (2007) and the references therein for details). However, none of the above papers deal with infinite horizon inventory models with fixed ordering cost. Our paper takes into account both risk aversion and ambiguity aversion preferences of an inventory planner in a unified framework and establishes the structure of the optimal policy for the infinite horizon model.

Our model can be regarded as an extension of Chen and Simchi-Levi (2004a) and Chen and Simchi-Levi (2004b), which analyze risk neutral finite and infinite horizon inventory and pricing models. It also extends to Chen et al. (2007), which focuses on finite horizon risk aversion inventory and pricing models. Compared with traditional risk neutral inventory models, the Bellman equation of our risk and ambiguity averse model essentially replaces the expectation operator with the so-called general certainty equivalent operator  $\mathcal{G}_{\Theta}^R(\cdot)$  (to be defined later in this paper). The analysis for the infinite horizon models, however, is much more involved. The main difficulty comes from the fact that the general certainty equivalent operator  $\mathcal{G}_{\Theta}^R(\cdot)$  is not additive. That is, unlike the expectation operator, for two uncertain values  $\tilde{\xi}$  and  $\tilde{\psi}$ , in general,  $\mathcal{G}_{\Theta}^R(\tilde{\xi} + \tilde{\psi}) \neq \mathcal{G}_{\Theta}^R(\tilde{\xi}) + \mathcal{G}_{\Theta}^R(\tilde{\psi})$ . As a result, the existing techniques to prove the optimality of a stationary  $(s, S, p)$  policy for the infinite horizon risk neutral model may not apply. For example, in the risk neutral case, it is possible to define the expected discounted or long run average profit for a given stationary  $(s, S, p)$  policy using the elementary renewal reward theory (see Chen and Simchi-Levi (2004b)). In our model, however, it is not clear how the elementary renewal reward theory, built upon the additivity property of the expectation operator, can be applied to define the value associated with a given stationary  $(s, S, p)$  policy. Similarly, the proof technique proposed in Huh and Janakiraman (2008) for the discounted infinite horizon inventory and pricing model, again built upon the additivity property of the expectation operator, may not be extended to analyze our model. Thus, we have to take a different approach. Even though our approach bears some similarities with Chen and Simchi-Levi (2004b), the logic is different and the analysis is more complicated.

The organization of this paper is as follows. In section 2, we introduce the finite and infinite horizon inventory and pricing models with risk and ambiguity aversion. We then prove the optimal structural policies for the infinite horizon model under both the discounted and the long run average criteria in section 3 and provide some concluding remarks in section 4.

**2. Model formulations.** In this section we first introduce the inventory and pricing system. We then provide the risk and ambiguity averse modeling framework in a finite horizon. Finally, we discuss in detail the infinite horizon ambiguity and risk averse inventory and pricing models. For an up-to-date review of the inventory and pricing literature, we refer to Chen and Simchi-Levi (2010).

**2.1. Joint inventory and pricing system.** Consider an inventory system where a decision maker makes replenishment and pricing decisions over  $T$  time periods. Here we denote  $t = 1$  to be the end of the planning horizon; therefore,  $t = \tau$  represents the  $(\tau - 1)$ th period to the end.

For each period  $t$ , let  $\tilde{d}_t$  denote the demand in period  $t$ . Demands in different periods are assumed to be independent. The decision maker may affect the period  $t$  demand through setting a selling price  $p_t$  for this period. We assume  $p_t$  is bounded, with lower and upper bounds  $\underline{p}_t$  and  $\bar{p}_t$ , respectively. To simplify notation, use  $\mathcal{P}_t$  to denote the interval  $[\underline{p}_t, \bar{p}_t]$ . Notice that when  $\underline{p}_t = \bar{p}_t$  for each period  $t$ , price is not

a decision variable and the problem is reduced to the traditional inventory control problem.

Throughout this paper, we concentrate on demand functions of the following form:  
 ASSUMPTION 2.1. *For  $t = 1, 2, \dots$ , the demand function satisfies*

$$\tilde{d}_t = D_t(p_t, \tilde{\varepsilon}_t) := \tilde{\beta}_t - \tilde{\alpha}_t p_t,$$

where  $\tilde{\varepsilon}_t = (\tilde{\alpha}_t, \tilde{\beta}_t)$ , and  $\tilde{\alpha}_t, \tilde{\beta}_t$  are nonnegative and represent the uncertainties in period  $t$ . Furthermore, we assume that there exists a constant  $\Xi > 0$  such that  $D(p, \tilde{\varepsilon}) \leq \Xi$  for any feasible  $p$  and any realization of  $\tilde{\varepsilon}$ .

Let  $x_t$  be the inventory level at the beginning of period  $t$ , and  $y_t$  the inventory order-up-to level. Lead time is assumed to be zero. The ordering cost function includes both a time independent fixed cost  $k$  and a possibly time dependent variable cost  $c_t$ . Therefore, the period  $t$  ordering cost is  $k\delta(y_t - x_t) + c_t(y_t - x_t)$ , where  $\delta(x)$  equals 1 if  $x > 0$  and 0 otherwise. We assume that unsatisfied demand is backlogged. As customary in most stochastic inventory models, we assume the inventory holding or backlogging cost function,  $h_t(x)$ , is convex in terms of  $x$ —the inventory level at the end of time period  $t$ .

At the beginning of period  $t$ , the inventory manager decides the order-up-to level  $y_t$  and the price  $p_t$ . Thus, given the initial inventory level  $x_t$ , the order-up-to level  $y_t$  and the realization of the uncertainty  $\tilde{\varepsilon}_t$ , the profit at period  $t$  is

$$\bar{P}_t(x_t, y_t, p_t; \tilde{\varepsilon}_t) = -k\delta(y_t - x_t) - c_t(y_t - x_t) + p_t D_t(p_t, \tilde{\varepsilon}_t) - h_t(y_t - D_t(p_t, \tilde{\varepsilon}_t)).$$

Finally, at the end of the planning horizon, we define  $\bar{P}_0(x_0, y_0, p_0; \tilde{\varepsilon}_0) = c_0 x_0$ , which implies that any inventory left has a unit salvage value  $c_0$  while any unsatisfied demand incurs a unit penalty cost  $c_0$ .

**2.2. Finite horizon risk and ambiguity averse model.** In this subsection, we introduce the finite horizon risk and ambiguity averse model, which lays the foundation for the infinite horizon model. Following Nilim and El Ghaoui (2005) and Iyengar (2005), we assume that the decision maker is only aware of a set to which the probability distribution of the uncertainty  $\tilde{\varepsilon}_t$  belongs, and model ambiguity aversion as a game between the decision maker and nature. That is, the decision maker maximizes her expected utility, with respect to probability distributions chosen by nature, who is an adversary choosing the probability distribution from an ambiguity set  $\Theta_t$  of probability distributions against the decision maker's objective.<sup>2</sup>

Let  $u_t(f_t)$  be the decision maker's utility function on consumption  $f_t$  in period  $t$ . Throughout this paper, we assume that the decision maker's utility function at period  $t$  is an exponential utility function:  $u_t(f_t) = -a_t e^{-f_t/\rho_t}$  with  $a_t \geq 0$  and the "risk tolerance parameter"  $\rho_t > 0$ . Given each period's initial wealth level  $w_t$ , and the risk free saving and borrowing interest rate  $r_f$ , or, equivalently, the discount rate

<sup>2</sup>The "recursive multiple prior" model (Epstein and Schneider (2003)) forms the decision theory foundation of our model. Following the recursive multiple priors setting, our model can be interpreted as allowing the decision maker to start from a set of priors and conduct appropriate Bayesian updating over time. The ambiguity set  $\Theta_t$  captures the updated set of priors up to period  $t$ . In our infinite horizon model, we focus on stationary ambiguity set  $\Theta$ . This does not mean there is no learning as time evolves. It simply implies that such a learning procedure does not bring more information (or shrink the set of priors) beyond certain periods, as illustrated in Epstein and Schneider (2007)—with recursive multiple priors, "...ambiguity need not vanish in the long run.... Instead, the agent moves towards a state of time-invariant ambiguity."

$\gamma = 1/(1 + r_f)$ , the consumption flow  $f_t$  is determined by

$$(2.1) \quad f_t = w_t - \gamma w_{t-1} + \bar{P}_t(x_t, y_t, p_t; \tilde{\varepsilon}_t).$$

We assume the “rectangularity property” (Epstein and Schneider (2003), Nilim and El Ghaoui (2005), and Iyengar (2005)), which essentially states that the set  $\Theta_t$  does not vary with previous realizations of demand or actions. Following Theorem 2.2 of Iyengar (2005), a risk and ambiguity averse inventory planner with additive utility over time solves the following dynamic programming recursion:

$$(2.2) \quad U_t(x_t, w_t) = \max_{y_t, p_t: y_t \geq x_t, p_t \in \mathcal{P}_t} \min_{\tilde{f}_{\tilde{\varepsilon}_t} \in \Theta_t} \mathbf{E}_{\tilde{f}_{\tilde{\varepsilon}_t}} \left[ \max_{w_{t-1}} \left\{ u_t(f_t) + U_{t-1}(y - D_t(p_t, \tilde{\varepsilon}_t), w_{t-1}) \right\} \right],$$

with boundary condition  $U_0(x_0, w_0) = u_0(w_0 + c_0 x_0)$ , where  $\tilde{f}_{\tilde{\varepsilon}_t}$  denotes a probability distribution in the ambiguity set  $\Theta_t$ .

To present the main dynamic program recursion, we first introduce the *general certainty equivalent operator*  $\mathcal{G}_{\Theta}^R(g(\tilde{\xi}))$  defined on a function  $g(\cdot)$  of an ambiguous uncertainty  $\tilde{\xi}$  as follows:

$$(2.3) \quad \mathcal{G}_{\Theta}^R(g(\tilde{\xi})) = \min_{\tilde{f}_{\tilde{\xi}} \in \Theta} -R \ln \mathbf{E}_{\tilde{f}_{\tilde{\xi}}} \left[ \exp \left\{ -\frac{1}{R} [g(\tilde{\xi})] \right\} \right].$$

The parameter  $R$  presents the risk aversion level while the set  $\Theta$  represents the ambiguity set of probability distributions. The operator  $\mathcal{G}_{\Theta}^R$  generalizes the certainty equivalent operator  $\mathcal{CE}_{\tilde{\xi}}^R$  defined in Chen et al. (2007) in the sense that when  $\Theta$  is a singleton taking distribution  $f_{\tilde{\xi}}$ ,  $\mathcal{G}_{\Theta}^R$  reduces to  $\mathcal{CE}_{\tilde{\xi}}^R$ .

We will not specify any particular ambiguity sets, other than that they satisfy certain technical conditions so that the minimization in the general certainty equivalent operator can always be attained. For instance, if we assume that the uncertainty  $\tilde{\xi}$  has a bounded support,  $g$  is continuous and the ambiguity set  $\Theta$  is compact in an appropriately defined function space (say,  $\mathcal{L}^2$  space for continuous uncertainties and  $\ell^2$  space for discrete uncertainties), then the minimization in (2.3) can be attained. Indeed, these conditions are satisfied for the models analyzed in this paper.

To facilitate future analysis, we also introduce the modified single period profit function excluding the fixed cost

$$P_t(y, p; \tilde{\varepsilon}_t) = (p - c_t) D_t(p, \tilde{\varepsilon}_t) - \hat{h}_t(y - D_t(p, \tilde{\varepsilon}_t)),$$

in which

$$(2.4) \quad \hat{h}_t(x) = (c_t - \gamma c_{t-1}) x + h_t(x)$$

is the modified inventory holding and backlogging cost function.

**PROPOSITION 2.1.** *The inventory and pricing decisions in the ambiguity and risk averse model (2.2) can be calculated through the following dynamic programming recursion:*

$$(2.5) \quad G_t(x) = \max_{y, p: y \geq x, p \in \mathcal{P}_t} -k\delta(y - x) + \mathcal{G}_{\Theta_t}^{R_t} [P_t(y, p; \tilde{\varepsilon}_t) + \gamma G_{t-1}(y - D_t(p, \tilde{\varepsilon}_t))],$$

with boundary condition  $G_0(x) = 0$ , and the effective risk tolerance

$$(2.6) \quad R_t = \sum_{\tau=0}^t \gamma^{t-\tau} \rho_{\tau}.$$

The derivation of the above result is parallel to that of Theorem 3.3 in Chen et al. (2007) and is omitted in this paper. Similar to the risk averse model, here  $G_t(x)$  can be considered as the certainty equivalent, referred to as the *general certainty equivalent*, of the consumption flow generated by running the inventory system starting from period  $t$  with an initial inventory level  $x$  to the end of the planning horizon.

In Online Appendix C (Chen and Sun (2011)), we demonstrate that for the finite horizon models, an  $(s, S, A, p)$  policy is optimal, in general, and provide conditions under which it reduces to an  $(s, S, p)$  policy. We also prove that the inventory control parameters are uniformly bounded.

**2.3. Infinite planning horizon.** Now we consider the infinite horizon inventory and pricing model with stationary model parameters. We study both the discounted and long run average criteria. As will be seen, even though their dynamic programming recursions are similar, the discounted and average profit cases come from quite different sources.

The discounted infinite horizon model is a natural extension to the finite horizon model with  $T \rightarrow \infty$ . The objective of the decision maker is defined as (2.2), subject to (2.1) with  $\gamma \in (0, 1)$  as  $t$  approaches infinity. Dropping the subscript  $t$  due to stationarity, the infinite horizon version of the dynamic programming recursion (2.5), also referred to as the *Bellman equation*, becomes

$$(2.7) \quad G(x) = \max_{y, p: y \geq x, p \in \mathcal{P}} -k\delta(y - x) + \mathcal{G}_{\Theta}^{R_\infty} [P(y, p; \tilde{\varepsilon}) + \gamma G(y - D(p; \tilde{\varepsilon}))] .$$

As  $t \rightarrow \infty$ , the effective risk tolerance becomes  $R_\infty = \rho/(1 - \gamma)$ .

Next we consider the long run average case of the ambiguity and risk averse inventory and pricing problem. One possible objective for the long run average case is based on the consumption model by maximizing the long run average of the general certainty equivalent of cash flows generated from the inventory system, i.e.,  $\liminf_{t \rightarrow \infty} G_t(x)/t$ , in which  $G_t$  is from the dynamic program (2.5). As the discounted factor  $\gamma \rightarrow 1$ , however, the effective risk tolerance factor  $R_\infty = \rho/(1 - \gamma)$  approaches infinity, as if the decision maker becomes risk neutral. Therefore, we follow an alternative objective, motivated by the long run average risk averse Markovian decision models analyzed in the robust control literature (see, for instance, Di Masi and Stettner (1999)). Specifically, we consider an ambiguity and risk averse long run average inventory and pricing problem for a given risk tolerance factor  $R$  modeled as the following maximization problem:

$$(2.8) \quad \max_{\omega \in \Omega} \min_{\tilde{f}_\varepsilon \in \Theta^\infty} \liminf_{T \rightarrow \infty} \frac{1}{T} \mathcal{CE}_\varepsilon^R [\mathbf{P}(T, \omega, \varepsilon)],$$

where  $\mathbf{P}(T, \omega, \varepsilon)$  is the total profit generated from the inventory system over a  $T$ -period planning horizon given the uncertainty  $\varepsilon = (\tilde{\varepsilon}_t)_{t=1}^\infty$  and inventory and pricing policy  $\omega$ , and  $\Theta^\infty$  is the infinite direct product of the ambiguity set  $\Theta$ . It is not hard to show that this maximization problem is equivalent to maximizing

$$\liminf_{t \rightarrow \infty} \frac{1}{t} G_t(x)$$

with the general certainty equivalent  $G_t$  defined as

$$(2.9) \quad G_t(x) = \max_{y, p: y \geq x, p \in \mathcal{P}} -k\delta(y - x) + \mathcal{G}_{\Theta}^R [P(y, p; \tilde{\varepsilon}) + G_{t-1}(y - D(p, \tilde{\varepsilon}))]$$

with  $G_0(x) = 0$ .

Although the above long run average case and the discounted case are derived from different origins, the similarity in their corresponding dynamic programming recursions allows us to introduce a unified Bellman equation,

$$(2.10) \quad \phi(x) = \max_{y: y \geq x} -k\delta(y-x) + \max_{p: p \in \mathcal{P}} \mathcal{G}_{\Theta}^R [P(y, p; \tilde{\varepsilon}) - \lambda + \gamma\phi(y - D(p, \tilde{\varepsilon}))].$$

In fact, when  $\gamma \in (0, 1)$ , the Bellman equation (2.7) for the discounted case can be written in the form (2.10) by simply setting  $R = R_{\infty}$  and

$$(2.11) \quad \phi(x) = G(x) - \frac{\lambda}{1-\gamma}.$$

Before proceeding to the proof for the optimal structural policies, we present a series of properties of the general certainty equivalent operator  $\mathcal{G}_{\Theta}^R$ , which will be useful in later proofs.

LEMMA 2.2. (a) Monotonicity: *If  $g(\tilde{\xi})$  is greater than  $h(\tilde{\xi})$  almost everywhere, then  $\mathcal{G}_{\Theta}^R(g(\tilde{\xi})) \geq \mathcal{G}_{\Theta}^R(h(\tilde{\xi}))$ .*

(b)  $\Delta$  property: *For any constant  $\Delta$ ,  $\mathcal{G}_{\Theta}^R(g(\tilde{\xi}) + \Delta) = \mathcal{G}_{\Theta}^R(g(\tilde{\xi})) + \Delta$ .*

(c) Contraction: *For any functions  $g$  and  $h$ ,  $|\mathcal{G}_{\Theta}^R(g(\tilde{\xi})) - \mathcal{G}_{\Theta}^R(h(\tilde{\xi}))| \leq \sup_{\tilde{\xi}} |g(\tilde{\xi}) - h(\tilde{\xi})|$ .*

(d) Preservation of concavity: *If  $g(x, \tilde{\xi})$  is concave,  $k$ -concave, or symmetric  $k$ -concave in  $x$  for any fixed  $\tilde{\xi}$ , then  $\mathcal{G}_{\Theta}^R(g(x, \tilde{\xi}))$  is concave,  $k$ -concave, or symmetric  $k$ -concave, respectively (see Online Appendix A (Chen and Sun (2011)) for these concepts and their properties).*

*Proof.* The proof for parts (a), (b), and (c) is straightforward and thus is omitted. Part (d) follows from Propositions A.8 and A.9 in Online Appendix A (Chen and Sun (2011)).  $\square$

The following lemma lists  $\mathcal{G}_{\Theta}^R(g(\tilde{\xi}))$ 's properties with respect to the risk aversion level  $R$  (see Online Appendix B (Chen and Sun (2011)) for their proof).

LEMMA 2.3. (a) Monotonicity in  $R$ : *For any  $0 < R' < R$ ,  $\mathcal{G}_{\Theta}^{R'}(g(\tilde{\xi})) \leq \mathcal{G}_{\Theta}^R(g(\tilde{\xi}))$ .*

(b) Concavity:  *$\mathcal{G}_{\Theta}^R(g(\tilde{\xi}))$  is concave in  $R$ .*

(c) Local boundedness of superderivative: *For any two constants  $\delta > 0$  and  $M > 0$ , there exists a constant  $\kappa > 0$  such that for any continuous function  $g(\tilde{\xi})$  with  $|g(\tilde{\xi})| \leq M$  for any  $\tilde{\xi}$ ,*

$$\mathcal{G}_{\Theta}^{R'}(g(\tilde{\xi})) - \mathcal{G}_{\Theta}^R(g(\tilde{\xi})) \leq \kappa(R' - R) \quad \text{for } R' \geq R \geq \delta.$$

**3. Structural policies for infinite horizon models.** In this section, we prove that  $(s, S, p)$  policies are optimal for the infinite horizon ambiguity and risk averse models.

THEOREM 3.1. *A stationary  $(s, S, p)$  policy is optimal for the infinite horizon stationary ambiguity and risk averse inventory and pricing models with either discounted or long run average optimization criteria.*

The proof of the theorem is quite involved and divided into several steps. First, we construct a solution for the Bellman equation (2.10). Then, we illustrate that the constructed solution is indeed optimal for the total discounted case and the long run average case, respectively.

**3.1. Bellman equation.** In this subsection, we analyze the Bellman equation (2.10). For simplicity, we drop the subscript and superscript in the general certainty equivalent operator  $\mathcal{G}_\Theta^R$ .

Assume that function

$$Q(x) = \max_{p \in \mathcal{P}} \mathcal{G}(P(x, p; \tilde{\varepsilon}))$$

is well defined and

$$\lim_{|x| \rightarrow \infty} Q(x) = -\infty.$$

Lemma 2.2 (d) implies that  $\mathcal{G}(P(x, p; \tilde{\varepsilon}))$  is jointly concave in  $x$  and  $p$  and hence  $Q(x)$  is concave in  $x$ .

To simplify our analysis, we assume that the realized demand is bounded below by a positive constant  $\eta > 0$ . That is,  $D(p, \tilde{\varepsilon}) \geq \eta$  for any feasible  $p$  and any realization of  $\tilde{\varepsilon}$ .<sup>3</sup>

To prove that a stationary  $(s, S)$  inventory policy is optimal, we define the following function in a recursive manner. It will become clear later on that this function is closely related to the value function of a given  $(s, S)$  inventory policy, which leads to a solution to the Bellman equation (2.10),

$$(3.1) \quad \varphi(x, s) = \begin{cases} 0 & \text{for } x \leq s, \\ \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s)] - Q(s) & \text{for } x > s. \end{cases}$$

Denote  $x^*$  one of the maximizers of  $Q(x)$ . We now show that  $\varphi$  satisfies the following properties.

**PROPOSITION 3.2.** (a)  $\varphi(x, s)$  is continuous in  $(x, s)$ . Therefore, the maximization in the formulation (3.1) is well defined.

(b) If  $s' \leq s \leq x^*$ , then  $\varphi(x, s') \geq \varphi(x, s)$  for any  $x$ . If, in addition,  $x \leq x^*$ , then  $\varphi(x, s) \geq 0$ .

(c) For any  $s \geq x^*$  and any  $x$ ,  $\varphi(x, s) \leq 0$ .

(d)  $\varphi(x, s) \rightarrow \infty$  for any fixed  $x$  as  $s \rightarrow -\infty$ .

(e) For any bounded set  $B$ ,  $\sup_{s \in B} \varphi(x, s) \rightarrow -\infty$  as  $x \rightarrow \infty$ .

(f) For any fixed  $s$ ,  $\varphi(x, s)$  is nondecreasing for  $x \leq y^*$ , where  $y^*$  is a minimizer of  $\hat{h}(x)$  defined in (2.4).

*Proof.* The basic idea is to use induction along  $x$ . That is, first assuming that the results hold for  $x \leq \bar{x}$ , we then show the results hold for  $x \in [\bar{x}, \bar{x} + \eta]$ .

The proof of part (a) is straightforward and thus is omitted.

We now prove part (b). First we show that  $\varphi(x, s) \geq 0$  for  $s \leq x^*$  and  $x \leq x^*$ . Observe that  $\varphi(x, s) = 0$  for any  $x \leq s$ . Now assume that for any  $x \leq \bar{x}$  for some  $\bar{x} \in [s, x^*]$ ,  $\varphi(x, s) \geq 0$ . Then for any  $x \in [\bar{x}, \min(\bar{x} + \eta, x^*)]$ , we have, following the induction hypothesis,

$$\varphi(x - D(p, \tilde{\varepsilon}), s) \geq 0,$$

which implies that

$$\varphi(x, s) = \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s)] - Q(s) \geq Q(x) - Q(s) \geq 0.$$

<sup>3</sup>Similar to the risk neutral inventory and pricing model analyzed in Chen and Simchi-Levi (2004b), this assumption can be relaxed by assuming that  $\min_{\tilde{\varepsilon} \in \Theta} \Pr(D(p, \tilde{\varepsilon}) = 0) \leq 1 - \kappa > 0$  for some  $1 > \kappa > 0$  and for any feasible  $p$ .

We now show that for  $s' \leq s \leq x^*$ ,  $\varphi(x, s') \geq \varphi(x, s)$  for any  $x$ . First, we show that this result holds for  $x \leq s$  by distinguishing the following cases.

- For any  $x \leq s'$ ,  $\varphi(x, s') = \varphi(x, s) = 0$ .
- For any  $x \in [s', s]$ ,  $x \leq x^*$  and thus  $\varphi(x - D(p, \tilde{\varepsilon}), s') \geq 0$ , which implies that for any realization of  $\tilde{\varepsilon}$ ,

$$P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s') \geq P(x, p; \tilde{\varepsilon}).$$

Therefore, for  $x \leq s \leq x^*$ ,

$$\varphi(x, s') \geq Q(x) - Q(s') \geq 0 = \varphi(x, s).$$

We now assume that  $\varphi(x, s') \geq \varphi(x, s)$  for  $x \leq \bar{x}$  for some  $\bar{x} \geq s$ . Then for any  $x \in [\bar{x}, \bar{x} + \eta]$ ,

$$\begin{aligned} \varphi(x, s') &= \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s')] - Q(s') \\ &\geq \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s)] - Q(s) \\ &= \varphi(x, s), \end{aligned}$$

where the inequality follows from Lemma 2.2 (a), the concavity of  $Q$ , and the fact that  $s' \leq s \leq x^*$  as well as the induction hypothesis. Therefore, the proof for part (b) is complete.

We now prove part (c). The induction hypothesis is that for  $s \geq x^*$ ,  $\varphi(x, s) \leq 0$  for any  $x \leq \bar{x}$ , with  $\bar{x} \geq s$ . The induction hypothesis holds by definition at  $\bar{x} = s$ . Then for  $x \in [\bar{x}, \bar{x} + \eta]$ , we have that  $\varphi(x - D(p, \tilde{\varepsilon}), s) \leq 0$  according to the induction hypothesis, which implies that

$$\begin{aligned} \varphi(x, s) &= \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s)] - Q(s) \\ &\leq \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon})] - Q(s) \\ &= Q(x) - Q(s) \leq 0, \end{aligned}$$

which completes the proof of part (c).

We now prove part (d). From part (b), we have that for any fixed  $x$  and  $s \leq x^*$ ,

$$\varphi(x - D(p, \tilde{\varepsilon}), s) \geq \varphi(x - D(p, \tilde{\varepsilon}), x^*).$$

Therefore,

$$\begin{aligned} \varphi(x, s) &= \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s)] - Q(s) \\ &\geq \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), x^*)] - Q(x^*) + Q(x^*) - Q(s) \\ &= \varphi(x, x^*) + Q(x^*) - Q(s) \rightarrow \infty \text{ as } s \rightarrow -\infty. \end{aligned}$$

We now prove part (e). First, we prove that for any  $x$ ,

$$(3.2) \quad \varphi(x, s) \leq l \left( \left\lceil \frac{x-s}{\eta} \right\rceil \right) (Q^* - Q(s)),$$

where  $Q^* = \max_x Q(x)$ ,  $\lceil x \rceil$  is the smallest integer no less than  $x$ , and function  $l$  is defined as follows for a nonnegative integer  $n$ :

$$l(n) = \begin{cases} \frac{1-\gamma^n}{1-\gamma} & \text{if } \gamma \in (0, 1), \\ n & \text{if } \gamma = 1. \end{cases}$$



Indeed, the inequality clearly holds for  $x \in (s, s + \eta]$ . Now assume that the inequality holds for  $x \leq \bar{x}$  for some  $\bar{x} > s$ . Then, for  $x \in (\bar{x}, \bar{x} + \eta]$ , we have that  $x - D(p, \tilde{\varepsilon}) \leq \bar{x}$  and

$$\begin{aligned} \varphi(x, s) &= \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s)] - Q(s) \\ &\leq \max_{p \in \mathcal{P}} \mathcal{G} \left[ P(x, p; \tilde{\varepsilon}) + \gamma l \left( \left\lceil \frac{x - \eta - s}{\eta} \right\rceil \right) (Q^* - Q(s)) \right] - Q(s) \\ &\leq Q^* - Q(s) + \gamma l \left( \left\lceil \frac{x - \eta - s}{\eta} \right\rceil \right) (Q^* - Q(s)) \\ &= l \left( \left\lceil \frac{x - s}{\eta} \right\rceil \right) (Q^* - Q(s)). \end{aligned}$$

Thus, (3.2) holds for any  $x$ . Since  $\lim_{|x| \rightarrow \infty} Q(x) = -\infty$ , for a given  $\sigma < Q(s)$ , there exists a constant  $x_\sigma$  such that  $Q(x) \leq \sigma$  for all  $x \geq x_\sigma$ . Similarly, we can prove that for any  $x \geq x_\sigma$ ,

$$\varphi(x, s) \leq l \left( \left\lceil \frac{x - x_\sigma}{\eta} \right\rceil \right) (\sigma - Q(s)) + \gamma^{\lceil \frac{x - x_\sigma}{\eta} \rceil} l \left( \left\lceil \frac{x_\sigma - s}{\eta} \right\rceil \right) (Q^* - Q(s)),$$

which implies that as  $x \rightarrow \infty$ ,  $\varphi(x, s) \rightarrow -\infty$  uniformly for  $s$  in any bounded set  $B$ .

Finally, we prove part (f). The induction hypothesis is that  $\varphi(x, s)$  is nondecreasing for  $x \leq \bar{x}$  for some  $\bar{x} \in [s, y^*]$ . We now focus on the case with  $x \in (\bar{x}, \min(\bar{x} + \eta, y^*)]$ . In this case, since  $\hat{h}$  is convex, for a given  $p$ ,  $\hat{h}(x - D(p, \tilde{\varepsilon}))$  is nonincreasing in  $x \leq y^*$ . In addition, the induction hypothesis implies that  $\varphi(x - D(p, \tilde{\varepsilon}), s)$  is nondecreasing in  $x$  for a given  $p$ . Thus, the function

$$\varphi(x, s) = \max_{p \in \mathcal{P}} \mathcal{G}[(p - c)D(p, \tilde{\varepsilon}) - \hat{h}(x - D(p, \tilde{\varepsilon})) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s)] - Q(s)$$

is nondecreasing for  $x \in (\bar{x}, \min(\bar{x} + \eta, y^*)]$ , and therefore for all  $x \leq y^*$ .  $\square$

Based on the above properties of function  $\varphi$ , we have the following result.

LEMMA 3.3. *If a function  $\varphi(x, s)$  satisfies Proposition 3.2, for any constant  $\kappa \geq 0$ , there exists  $s(\kappa) \leq x^*$  such that  $\max_x \varphi(x, s(\kappa)) = \kappa$ .*

*Proof.* Define  $f(s) = \max_x \varphi(x, s)$ . Since  $\varphi(x, s)$  is continuous in  $x$  and  $\varphi(x, s) \rightarrow -\infty$  as  $x \rightarrow \infty$ ,  $f(s)$  is well defined. In addition, the proof of Proposition 3.2 (c), together with the continuity of  $\varphi(x, s)$  in  $(x, s)$  implies that the  $f(s)$  is continuous. Finally, Proposition 3.2 (d) implies that  $\lim_{s \rightarrow -\infty} f(s) = \infty$ , while Proposition 3.2 (c) implies that  $f(s) \leq 0$  for  $s \geq x^*$ . Thus, there exists a constant  $s(\kappa) \leq x^*$  such that  $\kappa = f(s(\kappa)) = \max_x \varphi(x, s(\kappa))$ .  $\square$

In particular, denote  $s^* = s(k)$ , i.e.,  $\varphi(x, s^*) \leq \max_x \varphi(x, s^*) = k$ . Abuse notations and define  $\varphi(x) = \varphi(x, s^*)$ . Define

$$S^* = \max\{x : \varphi(x, s^*) = k\}.$$

We are now ready to construct a solution to Bellman equation (2.10) and show that the  $(s^*, S^*)$  policy define by

$$(3.3) \quad y^*(x) = \begin{cases} S^*, & x \leq s^*, \\ x, & \text{o.w.} \end{cases}$$

is optimal for (2.10).

**THEOREM 3.4.**  $(\phi(x) = \varphi(x), \lambda = Q(s^*))$  satisfies Bellman equation (2.10). Furthermore, the  $(s^*, S^*)$  policy solves the outer-maximization problem in (2.10).

The above result is similar to Theorem 5.1 of Chen and Simchi-Levi (2004b). Its proof contains several similar steps as well and is presented in the Online Appendix (Chen and Sun (2011)). However, in contrast to Chen and Simchi-Levi (2004b), which first characterize the best  $(s, S)$  inventory policy and show that its corresponding infinite horizon profit satisfies the Bellman equation, we start from (3.1), a recursion given by an  $(s, S)$  policy, and construct a solution to the Bellman equation (2.10). At this point, we have not shown that function  $\varphi(x)$  is a value function associated with the  $(s^*, S^*)$  policy. Furthermore, we have yet to show that the  $(s^*, S^*)$  inventory policy to the Bellman equation combined with the optimal pricing strategy  $p^*$  (the optimal solution for the inner-maximization problem in the Bellman equation (2.10) for a given  $y$ ) is optimal to the original problem. This is exactly what we are going to do in the following subsections. Before we proceed, we claim that for recursions (2.5) and (2.9), which define value function  $G_t(x)$  for the discounted case and the average case, respectively, there exists a constant  $\bar{S}$ , independent of  $t$ , such that it is optimal not to order for any  $x \geq \bar{S}$  (see Theorem C.3 in the Online Appendix (Chen and Sun (2011)) for the proof of this claim).

**3.2. Discounted case.** We prove the optimality of the stationary  $(s^*, S^*, p^*)$  policy in two steps: first, we show that  $G(x)$ , the solution to the Bellman equation (2.7) defined by

$$G(x) = \phi(x) + \frac{\lambda}{1-\gamma} = \varphi(x) + \frac{Q(s^*)}{1-\gamma},$$

is the infinite horizon value function of the stationary  $(s^*, S^*, p^*)$  policy; second, we show that the optimal finite horizon value function  $G_t(x)$  defined in the recursion (2.5) pointwise converges to  $G(x)$ . For this purpose, define finite horizon value iteration associated with the stationary policy  $(s^*, S^*, p^*)$ ,

$$(3.4) \quad \check{G}_t(x) = \begin{cases} -k + f_{\check{G}_{t-1}}^{R_t}(S^*), & x < s^*, \\ f_{\check{G}_{t-1}}^{R_t}(x), & x \geq s^*, \end{cases}$$

with boundary condition  $\check{G}_0(x) = 0$ , where

$$(3.5) \quad f_G^R(x) = \mathcal{G}_\Theta^R[P(x, p^*(x); \tilde{\varepsilon}) + \gamma G(x - D(p^*(x), \tilde{\varepsilon}))].$$

The validation of (3.4) follows a special case of Proposition 2.1, with the feasible set of inventory and pricing policies being a singleton.

Also from (2.11) and Theorem 3.4, we have that

$$G(x) = \begin{cases} -k + f_G^{R_\infty}(S^*), & x < s^*, \\ f_G^{R_\infty}(x), & x \geq s^*. \end{cases}$$

**THEOREM 3.5.** For any  $x$ , we have

- (a)  $\lim_{t \rightarrow \infty} \check{G}_t(x) = G(x)$  and
- (b)  $\lim_{t \rightarrow \infty} G_t(x) = G(x)$ .

*Proof.* (a) Pick an arbitrary constant  $\Delta$  with  $\Delta \geq S^*$ . For any  $x \leq \Delta$ ,  $y^*(x) \leq \Delta$ , where  $y^*(x)$  follows the  $(s^*, S^*)$  policy defined in (3.3). From the construction of  $G(x)$ , we can show that there exists a constant  $M$  such that

$$|P(y^*(x), p^*(y^*(x)); \tilde{\varepsilon}) + \gamma G(y^*(x) - D(p^*(y^*(x)), \tilde{\varepsilon}))| \leq M \text{ for any } x \leq \Delta.$$

Lemma 2.3 parts (a) and (c) imply that for any  $R_\infty \geq R_t \geq \rho > 0$  (recall that  $\rho$  is the risk aversion parameter),

$$(3.6) \quad 0 \leq f_G^{R_\infty}(y^*(x)) - f_G^{R_t}(y^*(x)) \leq \kappa(R_\infty - R_t) \text{ for any } x \leq \Delta$$

for some  $\kappa > 0$ . Therefore, for any  $x \leq \Delta$ ,

$$\begin{aligned} |\check{G}_t(x) - G(x)| &= \left| f_{\check{G}_{t-1}}^{R_t}(y^*(x)) - f_G^{R_\infty}(y^*(x)) \right| \\ &= \left| f_{\check{G}_{t-1}}^{R_t}(y^*(x)) - f_G^{R_t}(y^*(x)) + f_G^{R_t}(y^*(x)) - f_G^{R_\infty}(y^*(x)) \right| \\ &\leq \gamma \max_{z: z \leq \Delta} |\check{G}_{t-1}(z) - G(z)| + \kappa |R_t - R_\infty| \\ &\leq \gamma^t \max_{z: z \leq \Delta} |\check{G}_0(z) - G(z)| + \kappa \sum_{\tau=0}^t \gamma^{t-\tau} |R_\tau - R_\infty|, \end{aligned}$$

where the first inequality follows from Lemma 2.2 (c) and inequality (3.6), while the second inequality follows from using the first inequality repeatedly.

Since  $R_\infty - R_t = \frac{\gamma^{t+1}\rho}{1-\gamma}$ , we have that

$$\kappa \sum_{\tau=0}^t \gamma^{t-\tau} |R_\tau - R_\infty| = \frac{\rho\kappa}{1-\gamma} (t+1)\gamma^{t+1} \xrightarrow{t \rightarrow \infty} 0.$$

As the upper bound  $\Delta$  was chosen arbitrarily, part (a) is proved.

(b) For a given  $x$ , define  $y'$  and  $p'$  to be the optimal inventory and pricing decision in the Bellman equation (2.5) for  $G_t(x)$ . For any  $\Delta \geq \max(S^*, \bar{S})$  and  $x \leq \Delta$ , we have

$$G(x) - G_t(x) \leq \mathcal{G}_\Theta^{R_\infty}[P(y', p'; \tilde{\varepsilon}) + \gamma G(y - D(p', \tilde{\varepsilon}))] - \mathcal{G}_\Theta^{R_t}[P(y', p'; \tilde{\varepsilon}) + \gamma G_{t-1}(y - D(p', \tilde{\varepsilon}))],$$

which can be shown to approach 0 as  $t \rightarrow \infty$  following the same logic as in part (a) together with the fact that  $y'$  is bounded for all  $t$  for all  $x \leq \Delta$ .

Again if we define  $y''$  and  $p''$  to be the optimal inventory and pricing decision in the Bellman equation (3.6) for  $G(x)$ , we may repeat the above argument and show that  $G_t(x) - G(x)$  is bounded above by a term which approaches zero as  $t \rightarrow \infty$ .  $\square$

**3.3. Long run average case.** Again we show that the optimality of the stationary  $(s^*, S^*, p^*)$  policy in two steps. First, we prove that  $\lim_{t \rightarrow \infty} \frac{1}{t} G_t(x) = \lambda$ , in which  $\lambda = Q(s^*)$ , along with the function  $\phi(x)$ , constitutes the optimal solution to the Bellman equation (2.10) with  $\gamma = 1$ , and the function  $G_t(x)$  follows the finite horizon general certainty equivalent value iteration recursion (2.9). Second, we show that  $\lambda$  is indeed the long run average value associated with the stationary  $(s^*, S^*, p^*)$  policy, i.e.,  $\lim_{t \rightarrow \infty} \frac{1}{t} \check{G}_t(x) = \lambda$ , where  $\check{G}_t(x)$  is defined in (3.4) with  $\gamma = 1$ .

Recall Theorem 3.4,  $\lambda = Q(s^*)$ , and

$$\varphi(x) + \lambda = \begin{cases} -k + \mathcal{G}_\Theta^R[P(S^*, p^*(S^*); \tilde{\varepsilon}) + \varphi(S^* - D(p^*(S^*), \tilde{\varepsilon}))], & x < s^*, \\ \mathcal{G}_\Theta^R[P(x, p^*(x); \tilde{\varepsilon}) + \varphi(x - D(p^*(x), \tilde{\varepsilon}))], & x \geq s^*. \end{cases}$$

**THEOREM 3.6.** *Pointwise convergence is as follows:*

- (a)  $\lim_{t \rightarrow \infty} \frac{1}{t} \check{G}_t = \lambda$ .
- (b)  $\lim_{t \rightarrow \infty} \frac{1}{t} G_t = \lambda$ .

*Proof.* (a) We will show using induction that for any  $x \leq S^*$  and  $t$ , there is a bound  $M > 0$  such that

$$|\check{G}_t(x) - \varphi(x) - t\lambda| \leq M.$$

For period 0,  $\check{G}_0(x) = 0$  and  $0 \leq \varphi(x) \leq k$  for  $x \leq S^*$  (see Online Appendix Lemma D.1 parts (b) and (d)). The induction hypothesis is valid.

We have

$$\begin{aligned} & |\check{G}_t(x) - \varphi(x) - t\lambda| = |\check{G}_t(x) - (t-1)\lambda - (\varphi(x) + \lambda)| \\ &= \left| \mathcal{G}_\Theta^R[P(y^*(x), p^*(y^*(x))); \tilde{\varepsilon}] + \check{G}_{t-1}(y^*(x) - D(p^*(y^*(x)), \tilde{\varepsilon})) - (t-1)\lambda \right. \\ &\quad \left. - \mathcal{G}_\Theta^R[P(y^*(x), p^*(y^*(x))); \tilde{\varepsilon}] + \varphi(y^*(x) - D(p^*(y^*(x)), \tilde{\varepsilon})) \right| \\ &\quad (\text{definition of } y^*(x), \check{G}_t(x), \varphi(x)) \\ &\leq \max_{z: z \leq S^*} |\check{G}_{t-1}(z) - \varphi(z) - (t-1)\lambda| \quad (\text{Lemma 2.2 (c)}) \\ &\leq M \quad (\text{induction hypothesis}). \end{aligned}$$

Therefore

$$\lim_{t \rightarrow \infty} \frac{1}{t} \check{G}_t(x) - \lambda = 0 \quad \forall x \leq S^*.$$

For any given inventory level  $x > S^*$ , since the inventory level will decrease to no more than  $S^*$  in finite number of periods and the demand is lower bounded by  $\eta$ , the pointwise convergence result still holds.

(b) Observe that for any  $x \leq \Delta$  for some  $\Delta \geq \max(S^*, \bar{S})$ ,

$$\begin{aligned} & \varphi(x) + t\lambda - G_t(x) = (\varphi(x) + \lambda) + (t-1)\lambda - G_t(x) \\ &\leq \mathcal{G}_\Theta^R[P(y^*(x), p^*(y^*(x))); \tilde{\varepsilon}] + \varphi(y^*(x) - D(p^*(y^*(x)), \tilde{\varepsilon})) + (t-1)\lambda \\ &\quad - \mathcal{G}_\Theta^R[P(y^*(x), p^*(y^*(x))); \tilde{\varepsilon}] + G_{t-1}(y^*(x) - D(p^*(y^*(x)), \tilde{\varepsilon})) \\ &\leq \max_{z \leq \Delta} \{\varphi(z) + (t-1)\lambda - G_{t-1}(z)\}, \end{aligned}$$

where the first inequality follows from the definition of  $y^*(x)$ ,  $p^*(y)$ ,  $G_t(x)$ , and  $\varphi(x)$ . Following a similar induction argument as in (a) we can show that for  $x \leq \Delta$ ,  $\varphi(x) + t\lambda - G_t(x) \leq M$  for some constant  $M$ .

To show that  $G_t(x) - \varphi(x) - t\lambda \leq M$  for some constant  $M$  for  $x \leq \Delta$ , we denote  $y$  and  $p_t^*(y)$  to be the optimal inventory and pricing decisions for the  $G_t(x)$  value iteration in (2.9). Again, using induction we can show that for any  $x \leq \Delta$ ,

$$\begin{aligned} & G_t(x) - \varphi(x) - t\lambda = (G_t(x) - (t-1)\lambda) - (\varphi(x) + \lambda) \\ &\leq \mathcal{G}_\Theta^R[P(y, p_t^*(y)); \tilde{\varepsilon}] + G_{t-1}(y - D(p_t^*(y), \tilde{\varepsilon})) - (t-1)\lambda \\ &\quad - \mathcal{G}_\Theta^R[P(y, p_t^*(y)); \tilde{\varepsilon}] + \varphi(y - D(p_t^*(y), \tilde{\varepsilon})) \\ &\leq \max_{z \leq \Delta} \{G_{t-1}(z) - \varphi(z) - (t-1)\lambda\}, \end{aligned}$$

where the first inequality follows from the definition of  $y$ ,  $p_t^*(y)$ ,  $G_t(x)$ , and  $\varphi(x)$ .

The proof is now complete.  $\square$

**4. Concluding remarks.** In this paper, we show that for the infinite horizon ambiguity and risk averse inventory and pricing model, a stationary  $(s, S, p)$  policy is optimal under both the discounted and the long run average optimization criteria. This is interesting considering that a weaker structure,  $(s, S, A, p)$  policy, is optimal for their finite horizon counterparts.

Its proof is much more complicated than the corresponding one for the risk neutral model in Chen and Simchi-Levi (2004b). The difficulty comes from the nonlinearity of the general certainty equivalent operator  $\mathcal{G}_{\Theta}^R$ , which renders the techniques in Chen and Simchi-Levi (2004b) and Huh and Janakiraman (2008), built upon the linearity of the expectation operator, inapplicable. Chen and Simchi-Levi (2004b), for instance, derive the value function for a given  $(s, S)$  inventory policy associated with its optimal pricing strategy and illustrate that the value function of the best  $(s, S)$  inventory policy gives a solution to the Bellman equation and the original risk neutral problem. In this paper, on the other hand, we first construct a function defined in a recursive manner based on a reorder point  $s$ . We then illustrate that for an appropriately chosen  $s$ , this function satisfies the Bellman equation, and the optimal policy in the Bellman equation is given by the reorder point  $s$  together with an order-up-to level  $S$ . Next we show that this function is the value function of the stationary  $(s, S)$  inventory policy (together with its optimal pricing strategy), and is the highest achievable (therefore optimal) value function.

For the discounted infinite horizon inventory and pricing problem with risk neutrality and no ambiguity, Huh and Janakiraman (2008) propose an alternative approach and prove that as long as certain conditions hold, a stationary  $(s, S, p)$  policy is optimal. Their conditions are imposed on the expected single period profit and their argument relies heavily on the linearity of the expectation operator. It is an open question whether their approach can be extended to analyze the ambiguity and risk averse model here.

An interesting question is how the  $(s, S)$  parameters in the optimal inventory control policy vary with the risk and ambiguity aversion parameters. For example, does  $s$  or  $S$  change monotonically with the risk and ambiguity aversion parameters? Unfortunately, in general, the answer is negative. We constructed numerical examples which indicate that neither  $s$  nor  $S$  change monotonically with the risk tolerance parameter  $R$ . Numerical examples in Scarf (1958) implies monotonicity does not hold even in an ambiguity averse newsvendor setting.

**Acknowledgments.** We would like to thank Professor Sean Meyn for insightful discussions regarding risk averse Markovian decision processes. We also thank Professor Suresh Sethi, the associate editor, and two anonymous referees for constructive suggestions.

#### REFERENCES

- X. CHEN, M. SIM, D. SIMCHI-LEVI, AND P. SUN (2007), *Risk Aversion in Inventory Management*, *Oper. Res.*, 55, pp. 828–842.
- X. CHEN AND D. SIMCHI-LEVI (2004), *Coordinating Inventory Control and Pricing Strategies with Random Demand and Fixed Ordering Cost: The Finite Horizon Case*, *Oper. Res.*, 52, pp. 887–896.
- X. CHEN AND D. SIMCHI-LEVI (2004), *Coordinating Inventory Control and Pricing Strategies with Random Demand and Fixed Ordering Cost: The Infinite Horizon Case*, *Math. Oper. Res.*, 29, pp. 698–723.
- X. CHEN AND D. SIMCHI-LEVI (2010), *Pricing and inventory management*, in *Handbook on Pricing Management*, O. Ozer and B. Philips, eds., to appear.

- X. CHEN AND P. SUN (2011), *Optimal Structure Policies for Ambiguity and Risk Averse Inventory and Pricing Models: Online Appendix*; available online at <http://faculty.fuqua.duke.edu/~psun/bio/infinitehorizonApp.pdf>.
- G. B. DI MASI AND L. STETTNER (1999), *Risk-sensitive control of discrete-time Markov processes with infinite horizon*, *SIAM J. Control Optim.*, 38, pp. 61–78.
- L. G. EPSTEIN AND M. SCHNEIDER (2003), *Recursive multiple-priors*, *J. Econom. Theory*, 113, pp. 1–31.
- L. G. EPSTEIN AND M. SCHNEIDER (2007), *Learning under ambiguity*, *Rev. Econom. Stud.*, 74, pp. 1275–1303.
- X. GAN, S. SETHI, AND H. YAN (2004), *Coordination of supply chains with risk-averse agents*, *Prod. Oper. Manag.*, 13, pp. 135–149.
- W. T. HUH AND G. JANAKIRAMAN (2008), *(s, S) optimality in joint inventory-pricing control: An alternate approach*, *Oper. Res.*, 56, pp. 783–790.
- D. IGLEHART (1963), *Optimality of (s, S) policies in the infinite horizon dynamic inventory problem*, *Management Sci.*, 9, pp. 259–267.
- G. N. IYENGAR (2005), *Robust dynamic programming*, *Math. Oper. Res.*, 30, pp. 257–280.
- A. E. B. LIM AND J. G. SHANTHIKUMAR (2007), *Relative entropy, exponential utility, and robust dynamic pricing*, *Oper. Res.*, 55, pp. 198–214.
- A. NILIM AND L. EL GHAOU (2005), *Robust solutions to Markov decision problems with uncertain transition matrices*, *Oper. Res.*, 53, pp. 780–798.
- H. SCARF (1958), *A min-max solution to an inventory problem*, in *Studies in Mathematical Theory of Inventory and Production*, K. Arrow, S. Karlin, and H. Scarf, eds., Stanford University Press, Stanford, CA, Palo Alto, CA, pp. 201–209.
- M. SOBEL AND D. TURCIC (2008), *Risk Aversion and Supply Chain Contract Negotiation*, Technical report, Working paper, Department of Operations, Weatherhead School of Management, Case Western Reserve University, Cleveland, OH.
- Y. S. ZHENG (1991), *A simple proof for optimality of (s, S) policies in infinite-horizon inventory systems*, *J. Appl. Probab.*, 28, pp. 802–810.