Optimal structural policies for ambiguity and risk averse inventory and pricing models*

Xin Chen† Peng Sun‡

March 13, 2009

Abstract

This paper discusses multi-period 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 to establish the optimality of stationary (s, S, p) policies for the infinite horizon inventory and pricing models, for which the existing proof techniques for the risk neutral counterparts may not be extended.

*The research of the first author is supported by the National Science Foundation, Grant No. CMMI-0653909.

†Department of Industrial and Enterprise Systems Engineering University of Illinois at Urbana-Champaign, Urbana, IL 61801. Email: xinchen@uiuc.edu

‡Fuqua School of Business, Duke University Box 90120, Durham, NC 27708. Email: psun@duke.edu

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 planner 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 models with a fixed ordering cost and stochastic price-dependent demand. Our model and analysis extends Chen et al. (2007) to an infinite horizon, and with ambiguity aversion. The main contribution is to show that the optimal control policies for the proposed ambiguity and risk averse inventory and pricing models share similar structural properties as their conventional risk neutral counterparts in an infinite planning horizon (Chen and Simchi-Levi (2004b)). In particular, while an (s, S, A, p) policy is optimal for the finite horizon ambiguity and risk averse model, a stationary (s, S, p) policy is optimal for the infinite horizon model. (See Chen and Simchi-Levi (2004a) for the definition of the (s, S, A, p) and (s, S, p) policy structures.) We also include some additional new results for the finite horizon case in the appendix.

Our proof for the infinite horizon ambiguity and risk averse models, however, is much more involved compared with its risk neutral counterpart. The main difficulty comes from the fact that the general certainty equivalent operator $\mathcal{G}^R_\Theta(\cdot)$ (to be defined later) is not additive. That is, unlike the expectation operator, for two uncertain values $\tilde{\xi}$ and $\tilde{\psi}$, in general,

$$\mathcal{G}^R_\Theta(\tilde{\xi} + \tilde{\psi}) \neq \mathcal{G}^R_\Theta(\tilde{\xi}) + \mathcal{G}^R_\Theta(\tilde{\psi}) \ .$$

Therefore the existing techniques to prove the optimality of a stationary (s, S, p) policy (see, for example, Chen and Simchi-Levi (2004b) and Huh and Janakiraman (2004)) for the infinite horizon risk neutral model may not apply. Thus, although 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 model with risk and ambiguity aversion. Then we present the proof of the optimal structural policies for the infinite horizon mode with discounted as well as the long run average criteria in Section 3.

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 details the infinite horizon ambiguity and risk averse inventory and pricing models.

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 $t = 1, 2, \ldots$, the demand function satisfies*

$$\tilde{d}_t = D_t(p_t, \tilde{\varepsilon}_t) := \tilde{\beta}_t - \tilde{\alpha}_t p_t, \tag{1}$$

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

$$\delta(x) := \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

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 excluding the fixed cost 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. First consider the case in which there is no ambiguity in demand distribution. In this case, the risk averse inventory and pricing model is derived and analyzed in Chen et al. (2007). For the finite horizon model, we show, in Chen et al. (2007), that an (s, S) policy is optimal when demand is exogenous, i.e., price is not a decision variable; and an (s, S, A, p) policy is optimal when the price is a decision variable.

Further assume that the decision maker does not known the exact probability distribution for the uncertainty $\tilde{\varepsilon}_t$. 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 $\tilde{\varepsilon}_t$ 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 Θ_t of probability distributions against the decision maker's objective.

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

$$f_t = w_t - \gamma w_{t-1} + \bar{P}_t(x_t, y_t, p_t; \tilde{\varepsilon}_t) \ . \tag{2}$$

Assuming the "rectangularity property" (Epstein and Schneider (2003), Nilim and El Ghaoui (2005), and Iyengar (2005)), and following either Theorem 1 of Nilim and El Ghaoui (2005) or Theorem 2.2 of Iyengar (2005), a risk and ambiguity averse inventory planner with additive utility over time solves the following dynamic programming recursion,

$$U_t(x_t, w_t) = \max_{y_t, p_t: y_t \geq x_t, p_t \in \mathcal{P}_t} \min_{f_{\tilde{\varepsilon}_t} \in \Theta_t} \mathbf{E}_{f_{\tilde{\varepsilon}_t}} \left[\max_{w_{t-1}} \left\{ u_t(f_t) + U_{t-1}(y - D_t(p_t, \tilde{\varepsilon}_t), w_{t-1}) \right\} \right] , \qquad (3)$$

with boundary condition

$$U_0(x_0, w_0) = u_0(w_0 + c_0 x_0) \ . \qquad (4)$$

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

$$\mathcal{G}_\Theta^R\big(g(\tilde{\xi})\big) = \min_{f_{\tilde{\xi}} \in \Theta} -R \ln \mathbf{E}_{f_{\tilde{\xi}}} \left[\exp\left\{ -\frac{1}{R} \left[g(\tilde{\xi}) \right] \right\} \right] \ . \qquad (5)$$

The parameter R presents the risk aversion level while the set Θ 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 Θ 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 Θ is compact in an appropriately defined function space, say $\mathcal{L}^2$ space for continuous uncertainties and ℓ^2 space for discrete uncertainties, then the minimization in (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

$$P_t(y, p; \tilde{\varepsilon}_t) = (p - c_t)\, D_t(p, \tilde{\varepsilon}_t) - \hat{h}_t(y - D_t(p, \tilde{\varepsilon}_t)) \ , \qquad (6)$$

in which

$$\hat{h}_t(x) = (c_t - \gamma c_{t-1})\, x + h_t(x) \qquad (7)$$

is the modified inventory holding and backlogging cost function.

Proposition 1 *The inventory and pricing decisions in the ambiguity and risk averse model (3)-(4) can be calculated through the following dynamic programming recursion*

$$G_t(x) = \max_{y, p: y \geq x, p \in \mathcal{P}_t} -k\delta(y - x) + \mathcal{G}_{\Theta_t}^{R_t} \left[P_t(y, p; \tilde{\varepsilon}_t) + \gamma G_{t-1}(y - D_t(p, \tilde{\varepsilon}_t)) \right] \ , \qquad (8)$$

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

$$R_t = \sum_{\tau=0}^{t} \gamma^{t-\tau} \rho_\tau \ . \tag{9}$$

The proof is parallel to that of Theorem 3.3 in Chen et al. (2007), which is based on Smith (1998), 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 the Appendix G we present two theorems, which extend Chen et al. (2007) with structural results on finite horizon models.

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 $T \to \infty$. The objective of the decision maker is defined as (3), subject to (2) 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 (8), also referred to as the *Bellman equation*, becomes

$$G(x) = \max_{y,p: y \geq x, p \in \mathcal{P}} -k\delta(y - x) + \mathcal{G}_\Theta^{R_\infty} \left[P(y, p; \tilde{\varepsilon}) + \gamma G(y - D(p; \tilde{\varepsilon})) \right] \ . \tag{10}$$

As $t \to \infty$, the effective risk tolerance becomes

$$R_\infty = \frac{\rho}{1 - \gamma}. \tag{11}$$

We have not yet formally established any connection between the above Bellman equation with either the optimal value function of the original problem, or an optimal policy. In Section 3, we will construct a solution for problem (10), and show that this solution gives the optimal value function, and an (s, S, p) policy attaining the maximization in (10) is optimal for the inventory horizon inventory and pricing problem.

Next we consider the long run average case of the ambiguity and risk averse inventory and pricing problem. One possible choice of 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,

$$\liminf_{t\to\infty} \frac{1}{t} G_t \ ,$$

in which G_t is from the dynamic program (8). As the discounted factor $\gamma \to 1$, however, the effective risk tolerance factor $R_\infty = \rho/(1-\gamma)$ approaches infinity, as if the decision maker becomes risk neutral. We conjecture that in this case, our ambiguity and risk averse model, in the long run average infinite planning horizon framework, reduces to an ambiguity averse and risk neutral model. We leave this issue for future exploration and focus on 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)), an ambiguity and risk averse long run average inventory and pricing problem for a given risk tolerance factor R can be modelled as the following maximization problem:

$$\max_{\omega\in\Omega} \min_{f_\varepsilon\in\Theta^\infty} \liminf_{T\to\infty} \ \frac{1}{T} \mathcal{CE}_\varepsilon^R\big[\mathbf{P}(T,\omega,\varepsilon)\big] \ , \tag{12}$$

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 ω, and Θ^∞ is the infinite direct product of the ambiguity set Θ. It is not hard to show that this maximization problem is equivalent to maximizing

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

with the general certainty equivalent G_t defined as

$$G_t(x) = \max_{y,p:y\geq x,p\in\mathcal{P}} -k\delta(y-x) + \mathcal{G}_\Theta^R\left[P(y,p;\tilde{\varepsilon}) + G_{t-1}(y - D(p,\tilde{\varepsilon}))\right], \tag{13}$$

and $G_0(x) = 0$.

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

$$\phi(x) = \max_{y:y\geq x} -k\delta(y-x) + \max_{p:p\in\mathcal{P}} \mathcal{G}_\Theta^R\left[P(y,p;\tilde{\varepsilon}) - \lambda + \gamma\phi(y - D(p,\tilde{\varepsilon}))\right] \ . \tag{14}$$

In fact, when $\gamma \in (0,1)$, the Bellman equation (10) for the discounted case can be written in the form (14) by simply setting $R = R_\infty$ and

$$\phi(x) = G(x) - \frac{\lambda}{1-\gamma} \ . \tag{15}$$

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 1 *(a)* Monotonicity*: If $g(\tilde{\xi})$ is greater than $h(\tilde{\xi})$ almost everywhere, then*

$$\mathcal{G}_\Theta^R\Big(g(\tilde{\xi})\Big) \geq \mathcal{G}_\Theta^R\Big(h(\tilde{\xi})\Big) \ .$$

(b) Δ property*: For any constant Δ,*

$$\mathcal{G}_\Theta^R\Big(g(\tilde{\xi}) + \Delta\Big) = \mathcal{G}_\Theta^R\Big(g(\tilde{\xi})\Big) + \Delta \ .$$

(c) Contraction*: For any functions g and h,*

$$\left|\mathcal{G}_\Theta^R\Big(g(\tilde{\xi})\Big) - \mathcal{G}_\Theta^R\Big(h(\tilde{\xi})\Big)\right| \leq \sup_{\tilde{\xi}} \left|g(\tilde{\xi}) - h(\tilde{\xi})\right| \ .$$

(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 $\mathcal{G}_\Theta^R\Big(g(x, \tilde{\xi})\Big)$ is concave, k-concave or symmetric k-concave respectively (we present in Appendix A the definitions and properties of k-concavity and symmetric k-concavity).*

Proof. The proof for parts (a), (b) and (c) is straightforward and thus is omitted. Part (d) follows from Proposition 6 and Proposition 7 in Appendix A. ■

The following lemma lists $\mathcal{G}_\Theta^R\Big(g(\tilde{\xi})\Big)$'s properties with respect to the risk aversion level R.

Lemma 2 *(a)* Monotonicity in R*: For any $0 < R' < R$, $\mathcal{G}_\Theta^{R'}\Big(g(\tilde{\xi})\Big) \leq \mathcal{G}_\Theta^R\Big(g(\tilde{\xi})\Big)$.*

(b) Concavity*: $\mathcal{G}_\Theta^R\Big(g(\tilde{\xi})\Big)$ 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 $\tilde{\xi}$,*

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

Proof. See Appendix B. ■

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 (14). Then, we illustrate that the constructed solution is indeed optimal for the total discounted and the long run average case.

3.1 Bellman Equation

Recall the Bellman equation (14)

$$\phi(x) = \max_{y \geq x} -k\delta(y - x) + \max_{p \in \mathcal{P}} \mathcal{G}\left[P(y, p; \tilde{\varepsilon}) - \lambda + \gamma\phi(y - D(p, \tilde{\varepsilon}))\right] \ ,$$

where, for simplicity, we drop the subscript and superscript in the 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. In addition,

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

Lemma 1 (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 $\tilde{\varepsilon}$.[1]

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 (14).

$$\varphi(x, s) = \begin{cases} 0 \ , & \text{for } x \leq s, \\ \max_{p \in \mathcal{P}} \mathcal{G}[P(x, p; \tilde{\varepsilon}) + \gamma\varphi(x - D(p, \tilde{\varepsilon}), s)] - Q(s) \ , & \text{for } x > s. \end{cases} \tag{16}$$

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

[1]This lower bound assumption can be relaxed by assuming that $\min_{f_{\tilde{\varepsilon}} \in \Theta} \Pr(D(p, \tilde{\varepsilon}) = 0) \leq 1 - \kappa > 0$ for some $1 > \kappa > 0$ and for any feasible p. The proof goes through with this relaxation following an argument similar to the approach used for the risk neutral inventory and pricing model analyzed in Chen and Simchi-Levi (2004b) and hence is omitted here.

Proposition 2 *(a) $\varphi(x,s)$ is continuous in (x,s). Therefore, the maximization in the formulation (16) is well defined.*

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

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

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

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

(f) For any fixed s, $\varphi(x,s)$ is nondecreasing for $x \le y^*$, where y^* is a minimizer of $\hat{h}(x)$ defined in Eq. (7).

Proof. The basic idea is to use induction along x. That is, first assuming that the results hold for $x \le \bar{x}$, 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 $\varphi(x,s) \ge 0$ for $s \le x^*$ and $x \le x^*$. Observe that $\varphi(x,s) = 0$ for any $x \le s$. Now assume that for any $x \le \bar{x}$ for some $\bar{x} \in [s, x^*]$, $\varphi(x,s) \ge 0$. Then for any $x \in [\bar{x}, \min(\bar{x}+\eta, x^*)]$, we have, following the induction hypothesis,

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

which implies that

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

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

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

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

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

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

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

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

where the inequality follows from Lemma 1 (a), the concavity of Q and the fact that $s' \le s \le x^*$ 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 $s \ge x^*$, $\varphi(x,s) \le 0$ for any $x \le \bar{x}$, with $\bar{x} \ge s$. The induction hypothesis holds by definition at $\bar{x} = s$. Then for $x \in [\bar{x}, \bar{x}+\eta]$, we have that $\varphi(x - D(p,\tilde{\varepsilon}), s) \le 0$ according to the induction hypothesis, which implies that

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

which completes the proof of part (c).

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

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

Therefore,

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

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

$$\varphi(x,s) \le l\left(\left\lceil \frac{x-s}{\eta} \right\rceil\right)(Q^* - Q(s)), \tag{17}$$

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 $x \le \bar{x}$ for some $\bar{x} > s$. Then, for $x \in (\bar{x}, \bar{x}+\eta]$, we have that $x - D(p,\tilde{\varepsilon}) \le \bar{x}$ and

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

Thus, (17) holds for any x. Since $\lim_{|x|\to\infty} Q(x) = -\infty$, for a given $\sigma < Q(s)$, there exists a constant x_σ such that $Q(x) \le \sigma$ for all $x \ge x_\sigma$. Similarly, we can prove that for any $x \ge x_\sigma$,

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

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

Finally, we prove part (f). The induction hypothesis is that $\varphi(x,s)$ is nondecreasing for $x \le \bar{x}$ 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 non-increasing in $x \le y^*$. 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

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

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

Based on the above properties of function φ, we have the following result.

Lemma 3 *If a function $\varphi(x,s)$ satisfies Proposition 2, for any constant $\kappa \ge 0$, there exists $s(\kappa) \le x^*$ 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 $\varphi(x,s) \to -\infty$ as $x \to \infty$, $f(s)$ is well-defined. In addition, the proof of Proposition 2 (c), together with the continuity of $\varphi(x,s)$ in (x,s) implies that the $f(s)$ is continuous. Finally, Proposition 2 (d) implies that $\lim_{s\to-\infty} f(s) = \infty$, while Proposition 2 (c) implies that $f(s) \le 0$ for $s \ge x^*$. Thus, there exists a constant $s(\kappa) \le x^*$ such that $\kappa = f(s(\kappa)) = \max_x \varphi(x, s(\kappa))$. ■

In particular, denote $s^* = s(k)$, i.e.,

$$\varphi(x, s^*) \le \max_x \varphi(x, s^*) = k \ .$$

Abuse notations and define $\varphi(x) = \varphi(x, s^*)$. And denote

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

We now list several important properties of $\varphi(x)$.

Lemma 4 *(a) $\varphi(x) = 0$ for any $x \le s^*$.*

(b) $\varphi(x) \le \varphi(S^*) = k$ for any x.

(c) $Q(x) \geq Q(s^*)$ *for any* $x \in [s^*, S^*]$.

(d) $\varphi(x) \geq 0$ *for any* $x \leq S^*$.

(e) $s^* \leq x^*$.

(f) $S^* \geq y^*$.

Lemma 4 above is similar to Lemma 4 in Chen and Simchi-Levi (2004b), which characterizes the best (s, S) policy in the risk neutral joint inventory and pricing model. It is also essentially parallel to Lemma 1 in Zheng (1991) for the standard infinite horizon inventory control (without pricing) model as a characterization of the best (s, S) policy. We present the proof in Appendix C. A significant difference in our analysis, however, is that we have not yet proven that the (s^*, S^*) policy given in Lemma 4 is the best among all (s, S) policies.

Based on Lemma 4 and following the same logic as in Lemma 5 and Theorem 5.1 of Chen and Simchi-Levi (2004b), we have the following results, whose proofs are in Appendixes D and E.

Lemma 5 $\varphi(x)$ *is a symmetric* k*-concave function.*

Theorem 3.2 $(\phi(x) = \varphi(x), \lambda = Q(s^*))$ *satisfies Bellman equation (14). Furthermore, policy*

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

solves the maximization problem in Equation (14).

At this point we shall point out a significant difference in the above proof to the one in Chen and Simchi-Levi (2004b). Chen and Simchi-Levi (2004b) first characterizes the best (s, S) inventory policy and show that its corresponding infinite horizon profit satisfies the Bellman equation. Here, however, starting from Eq. (16), a recursion given by an (s, S) policy, we construct a solution to the Bellman equation (14). 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 solution to the Bellman equation is optimal to the original problem.

In the following subsections, we are going to show that the value function $\varphi(x)$ is indeed associated with the (s^*, S^*) policy and that the stationary (s^*, S^*, p^*) policy is optimal for the discounted case and average case respectively, where p^* the optimal solution for the inner maximization in the Bellman equation (14) for a given y.

3.2 Discounted case

So far we have constructed a solution to the Bellman equation (14) and shown that the (s^*, S^*) policy attains its outside maximization. However, since we have not formally established the connections between the Bellman equation and either the optimal value function or the value function associated with the (s^*, S^*) policy, we cannot directly claim that the stationary (s^*, S^*) inventory policy is optimal.

One might attempt to use the approach employed in Theorem 7.1 in Chen and Simchi-Levi (2004b), i.e., to show that $G_t(x)$ defined in (8) converges to $G(x)$, the solution to the Bellman Equation (10), or

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

However, even if we show that $G_t(x)$ converges to $G(x)$, we cannot immediately claim that the optimality of the stationary (s^*, S^*) inventory policy. In fact, in Theorem 7.1 of Chen and Simchi-Levi (2004b), it was shown that the finite horizon value function converges to the value function given by a stationary (s, S) policy associated with its optimal pricing strategies. While here, the relationship of the stationary (s^*, S^*) policy and the function $G(x)$ is not that obvious, i.e., we need to prove that $G(x)$ is the infinite horizon value function of the stationary (s^*, S^*) policy and its corresponding optimal pricing strategy p^*.

Therefore below we prove the optimality of the stationary (s^*, S^*, p^*) policy in two steps: first, show that $G(x)$ is the infinite horizon value function of the stationary (s^*, S^*, p^*) policy; second, show that the optimal finite horizon value function $G_t(x)$ point-wise converges to $G(x)$. For this purpose, define finite horizon value iteration according to stationary policy (s^*, S^*, p^*),

$$\check{G}_t(x) = \begin{cases} -k + f_t(S^*), & x < s^* \\ f_t(x), & x \geq s^* \end{cases} , \tag{18}$$

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

$$f_t(x) = \mathcal{G}_{\Theta}^{R_t}[P(x, p^*(x); \tilde{\varepsilon}) + \gamma \check{G}_{t-1}(x - D(p^*(x), \tilde{\varepsilon}))].$$

The validation of (18) follows a special case of Proposition 1, with the feasible set of inventory and pricing policies being a singleton.

Also recall equations (8), (15) and Theorem 3.2,

$$\begin{aligned} G(x) &= \max_{y,p: y \geq x, p \in \mathcal{P}} -k\delta(y-x) + \mathcal{G}_{\Theta}^{R_\infty}\left[P(y, p; \tilde{\varepsilon}) + \gamma G(y - D(p, \tilde{\varepsilon}))\right] \\ &= \begin{cases} -k + f^{R_\infty}(S^*), & x < s^* \\ f^{R_\infty}(x), & x \geq s^* \end{cases} , \end{aligned} \tag{19}$$

where function f^R for a risk tolerance parameter R is defined as

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

Theorem 3.3 *For any x,*

(a) $\lim_{t\to\infty} \check{G}_t(x) = G(x)$*, and*

(b) $\lim_{t\to\infty} G_t(x) = G(x)$*.*

Proof. (a) For any x, define

$$y(x) = \begin{cases} S^*, & \text{if } x \le s^* \\ x, & \text{if } x > s^*. \end{cases}$$

Pick an arbitrary constant Δ with $\Delta \ge \max(\bar{S}, S^*)$. Then for any $x \le \Delta$, $y(x) \le \Delta$. In addition, from the construction of $G(x)$, we can show that there exists a constant M such that

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

Lemma 2 parts (a) and (c) imply that for any $R_\infty \ge R_t \ge \rho > 0$ (recall that ρ is the risk aversion parameter),

$$0 \le f^{R_\infty}(y(x)) - f^{R_t}(y(x)) \le \kappa(R_\infty - R_t), \text{ for any } x \le \Delta \tag{20}$$

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

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

where the first inequality follows from Lemma 1(c) and inequality (20), and the second inequality follows from using the first inequality repeatedly.

Recall (9) and (11),

$$R_\infty - R_t = \frac{\gamma^{t+1}\rho}{1-\gamma} \ .$$

Therefore, we have for any $x \le \Delta$,

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

Thus,

$$\lim_{t\to\infty} \left|\check{G}_t(x) - G(x)\right| = 0 \ .$$

Since the upper bound Δ was chosen arbitrarily, we have the point-wise convergence result.

(b) For a given x, define y' and p' to be the optimal inventory and pricing decision in the Bellman equation (8) for $G_t(x)$. We have

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

which approached 0 with $t \to \infty$ following the same logic as in (a) together with the fact that y' is bounded for all t, which is guaranteed by Theorem G.3.

Again if we define y'' and p'' to be the optimal inventory and pricing decision in the Bellman equation (19) for $G(x)$, we may repeat the above argument and show that $G_t(x) - G(x) \to 0$.

■

3.3 Long run average case

Recall the finite horizon general certainty equivalent value iteration recursion (13) when the discount factor $\gamma = 1$ and the risk aversion factor R

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

Next, we are going to show that $\lim_{t\to\infty} \frac{1}{t} G_t(x) = \lambda$, in which λ, along with the function $\phi(x)$, constitutes the optimal solution to the Bellman equation (14) with $\gamma = 1$. Again, this result alone is not sufficient for the claim that the stationary (s^*, S^*, p^*) policy is optimal. We also need to show that λ is indeed the long run average value given by the stationary (s^*, S^*, p^*) policy, i.e., if we denote the general certainty equivalent value iteration following the stationary (s^*, S^*, p^*) policy

$$\check{G}_t(x) = \begin{cases} -k + \mathcal{G}_\Theta^R[P(S^*, p^*(S^*);\tilde{\varepsilon}) + \check{G}_{t-1}(S^* - D(p^*(S^*),\tilde{\varepsilon}))], & x < s^* \\ \mathcal{G}_\Theta^R[P(x, p^*(x);\tilde{\varepsilon}) + \check{G}_{t-1}(x - D(p^*(x),\tilde{\varepsilon}))], & x \ge s^* \end{cases} , \qquad (21)$$

with boundary condition $\check{G}_0(x) = 0$, we need to show $\lim_{t\to\infty} \frac{1}{t}\check{G}_t(x) = \lambda$.

Recall Theorem 3.2, $\lambda = Q(s^*)$ and

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

Theorem 3.4 *Point-wise convergence*

(a) $\lim_{t\to\infty} \frac{1}{t}\check{G}_t = \lambda$

(b) $\lim_{t\to\infty} \frac{1}{t} G_t = \lambda$

In the interest of saving space, we move the proof to Appendix F.

We also provide some concluding remarks and additional discussions in Appendix H.

Acknowledgement We would like to thank Professor Sean Meyn for insightful discussion regarding risk averse Markovian decision processes.

References

Chen, X., Sim, M., Simchi-levi, D., & Sun, P. (2007). Risk aversion in inventory management. *Operations Research*, *55*(5), 828-842.

Chen, X., & Simchi-Levi, D. (2004a). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research*, *52*(6), 887-896.

Chen, X., & Simchi-Levi, D. (2004b). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Mathematics of Operations Research*, *29*(3), 698-723.

Di Masi, G., & Stettner, L. (1999). Risk-sensitive control of discrete-time markov processes with infinite horizon. *SIAM J. Control Optim.*, *38*(1), 61-78.

Epstein, L., & Schneider, M. (2003). Recursive multiple-priors. *Journal of Economic Theory*, *113*(1), 1-31.

Huh, W. T., & Janakiraman, G. (2004). Alternative approaches to important results in inventory and pricing control. *Operations Research*. (To appear)

Iyengar, G. (2005). Robust dynamic programming. *Mathematics of Operations Research*, *30*(2), 257-280.

Nilim, A., & El Ghaoui, L. (2005). Robust solutions to markov decision problems with uncertain transition matrices. *Operations Research*, *53*(5), 780-798.

Smith, J. E. (1998). Evaluating income streams: A decision analysis approach. *Management Science*, *44*(12), 1690-1708.

Zheng, Y. S. (1991). A simple proof for optimality of (s, s) policies in infinite-horizon inventory systems. *J. App. Pro.*, *28*, 802-810.

Online Appendix

A Review on k-convexity and symmetric k-convexity

In this section, we review some important properties of k-convexity and symmetric k-convexity that are used in this paper; see Chen (2003) for more details.

The concept of k-convexity was introduced by Scarf (1960) to prove the optimality of an (s, S) for the traditional inventory control problem.

Definition A.1 *A real-valued function f is called k-convex for $k \geq 0$, if for any $x_0 \leq x_1$ and $\lambda \in [0, 1]$,*

$$f((1-\lambda)x_0 + \lambda x_1) \leq (1-\lambda)f(x_0) + \lambda f(x_1) + \lambda k. \tag{22}$$

Below we summarize properties of k-convex functions.

Lemma 6 *(a) A real-valued convex function is also 0-convex and hence k-convex for all $k \geq 0$. A k_1-convex function is also a k_2-convex function for $k_1 \leq k_2$.*

(b) If $f_1(y)$ and $f_2(y)$ are k_1-convex and k_2-convex respectively, then for $\alpha, \beta \geq 0$, $\alpha f_1(y) + \beta f_2(y)$ is $(\alpha k_1 + \beta k_2)$-convex.

(c) If $f(y)$ is k-convex and w is a random variable, then $E\{f(y-w)\}$ is also k-convex, provided $E\{|f(y-w)|\} < \infty$ for all y.

(d) Assume that f is a continuous k-convex function and $f(y) \to \infty$ as $|y| \to \infty$. Let S be a minimum point of g and s be any element of the set

$$\{x | x \leq S, f(x) = f(S) + k\}.$$

Then the following results hold.

(i) $f(S) + k = f(s) \leq f(y)$, for all $y \leq s$.

(ii) $f(y)$ is a non-increasing function on $(-\infty, s)$.

(iii) $f(y) \leq f(z) + k$ for all y, z with $s \leq y \leq z$.

Proposition 3 *If $f(x)$ is a K-convex function, then function*

$$g(x) = \min_{y \geq x} Q\delta(y-x) + f(y),$$

is $\max\{K, Q\}$-convex.

Recently a weaker concept of symmetric k-convexity was introduced by Chen and Simchi-Levi (2004a, 2004b) when they analyze the joint inventory and pricing problem with fixed ordering cost.

Definition A.2 *A function $f : \Re \to \Re$ is called symmetric k-convex for $k \geq 0$, if for any $x_0, x_1 \in \Re$ and $\lambda \in [0, 1]$,*

$$f((1-\lambda)x_0 + \lambda x_1) \leq (1-\lambda)f(x_0) + \lambda f(x_1) + \max\{\lambda, 1-\lambda\}k. \tag{23}$$

A function f is called symmetric k-concave if $-f$ is symmetric k-convex.

Observe that k-convexity is a special case of symmetric k-convexity. The following results describe properties of symmetric k-convex functions, properties that are parallel to those summarized in Lemma 6 and Proposition 3. Finally, notice that the concept of symmetric k-convexity can be easily extended to *multi-dimensional* space.

Lemma 7 *(a) A real-valued convex function is also symmetric 0-convex and hence symmetric k-convex for all $k \geq 0$. A symmetric k_1-convex function is also a symmetric k_2-convex function for $k_1 \leq k_2$.*

(b) If $g_1(y)$ and $g_2(y)$ are symmetric k_1-convex and symmetric k_2-convex respectively, then for $\alpha, \beta \geq 0$, $\alpha g_1(y) + \beta g_2(y)$ is symmetric $(\alpha k_1 + \beta k_2)$-convex.

(c) If $g(y)$ is symmetric k-convex and w is a random variable, then $E\{g(y-w)\}$ is also symmetric k-convex, provided $E\{|g(y-w)|\} < \infty$ for all y.

(d) Assume that g is a continuous symmetric k-convex function and $g(y) \to \infty$ as $|y| \to \infty$. Let S be a global minimizer of g and s be any element from the set

$$X := \{x | x \leq S, g(x) = g(S) + k \text{ and } g(x') \geq g(x) \text{ for any } x' \leq x\}.$$

Then we have the following results.

(i) $g(s) = g(S) + k$ and $g(y) \geq g(s)$ for all $y \leq s$.

(ii) $g(y) \leq g(z) + k$ for all y, z with $(s+S)/2 \leq y \leq z$.

Proposition 4 *If $f(x)$ is a symmetric K-convex function, then the function*

$$g(x) = \min_{y \leq x} Q\delta(x-y) + f(y)$$

is symmetric $\max\{K,Q\}$*-convex. Similarly, the function*

$$h(x) = \min_{y \geq x} Q\delta(x-y) + f(y)$$

is also symmetric $\max\{K,Q\}$*-convex.*

Proposition 5 *Let $f(\cdot,\cdot)$ be a function defined on $\Re^n \times \Re^m \to \Re$. Assume that for any $x \in \Re^n$ there is a corresponding set $C(x) \subset \Re^m$ such that the set $C \equiv \{(x,y) \mid y \in C(x), x \in \Re^n\}$ is convex in $\Re^n \times \Re^m$. If f is symmetric k-convex over the set C, and the function*

$$g(x) = \inf_{y \in C(x)} f(x,y)$$

is well defined, then g is symmetric k-convex over $\Re^n$.

Proposition 6 *(see Chen et al. (2007) and Simchi-Levi et al. (2004)) If $g(x,\tilde{\xi})$ is concave, k-concave or symmetric k-concave for any given $\tilde{\xi}$, then $\mathcal{CE}^R_{\tilde{\xi}}(g(x,\tilde{\xi}))$ is also concave, k-concave or symmetric k-concave.*

Proposition 7 *If $f(x,t)$ is concave, k-concave or symmetric k-concave for any given t. Then $g(x) = \min_t f(x,t)$ (assumed to be well-defined) is also concave, k-concave or symmetric k-concave.*

The proof for this proposition is straightforward and omitted.

B Proof of Lemma 2

(a) It is a well known result (Pratt (1964)) that the certainty equivalent is increasing with the risk tolerance level. That is, $\mathcal{CE}^{R'}_{\tilde{\xi}}\left(g(\tilde{\xi})\right) \leq \mathcal{CE}^{R}_{\tilde{\xi}}\left(g(\tilde{\xi})\right)$ for any $f_{\tilde{\xi}} \in \Theta$. Therefore

$$\mathcal{G}^{R'}_{\Theta}\left(g(\tilde{\xi})\right) = \min_{f_{\tilde{\xi}} \in \Theta} \mathcal{CE}^{R'}_{\tilde{\xi}}\left(g(\tilde{\xi})\right) \leq \min_{f_{\tilde{\xi}} \in \Theta} \mathcal{CE}^{R}_{\tilde{\xi}}\left(g(\tilde{\xi})\right) = \mathcal{G}^{R}_{\Theta}\left(g(\tilde{\xi})\right).$$

(b) Define $h(x) = \mathcal{G}^1_{f_{\tilde{\xi}}}\left(xg(\tilde{\xi})\right)$ and

$$H(x,R) = \begin{cases} Rh(x/R), & \text{if } R > 0, \\ 0, & \text{if } R = 0, x = 0, \\ +\infty, & \text{if } R < 0. \end{cases}$$

Lemma 1 part (d) implies that h is concave, which in turn implies that $H(x,R)$ is concave in (x,R) (see Rockafellar (1970), page 35). Thus, $\mathcal{G}^R_{f_{\tilde{\xi}}}\left(g(\tilde{\xi})\right) = H(1,R)$ is concave in R for $R > 0$.

(c) For a given $f_{\tilde{\xi}} \in \Theta$,

$$\frac{\partial \mathcal{CE}_{\tilde{\xi}}^R(g(\tilde{\xi}))}{\partial R} = -\ln \mathbf{E}\left[\exp\{-g(\tilde{\xi})/R\}\right] - \frac{\mathbf{E}\left[g(\tilde{\xi})\exp\{-g(\tilde{\xi})/R\}\right]}{R\mathbf{E}\left[\exp\{-g(\tilde{\xi})/R\}\right]}.$$

(It is easy to verify that under our assumptions, the integration and differential can be interchanged when taking the derivative of $\mathcal{CE}_{\tilde{\xi}}^R(g(\tilde{\xi}))$ with respect to R.) The above formula immediately implies that 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 $\tilde{\xi}$,

$$0 \leq \frac{\partial \mathcal{CE}_{\tilde{\xi}}^R(g(\tilde{\xi}))}{\partial R} \leq \kappa.$$

Since when $f_{\tilde{\xi}} \in \Theta$ attains the minimum in the definition of $\mathcal{G}_{\Theta}^R$, we have that for any $R' \geq R$,

$$\mathcal{G}_{\Theta}^{R'}\big(g(\tilde{\xi})\big) - \mathcal{G}_{\Theta}^{R}\big(g(\tilde{\xi})\big) \leq \mathcal{CE}_{\tilde{\xi}}^{R'}(g(\tilde{\xi})) - \mathcal{CE}_{\tilde{\xi}}^{R}(g(\tilde{\xi})) \leq \left|\frac{\partial \mathcal{CE}_{\tilde{\xi}}^R(g(\tilde{\xi}))}{\partial R}\right|(R' - R) \leq \kappa(R' - R).$$

Thus, part (c) holds.

C Proof for Lemma 4

Part (a) follows directly from the construction of function $\varphi(x) = \varphi(x, s^*)$. Part (b) follows from the definition of s^*, and part (e) follows from Lemma 3. From proposition 2 (f), we have $S^* \geq y^*$ and thus part (f) holds.

We now prove part (c). Note that $\varphi(S^* - D(p, \tilde{\varepsilon}), s^*) \leq \varphi(S^*, s^*) = k$. We have that

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

in which the last equation follows Lemma 1(b). Therefore,

$$Q(S^*) - Q(s^*) \geq (1-\gamma)\varphi(S^*, s^*) = (1-\gamma)k \geq 0.$$

Since $Q(x)$ is concave and $s^* \leq x^*$, we have $Q(x) \geq Q(s^*)$ for $x \in [s^*, S^*]$, i.e., part (c) holds.

We now prove part (d). Clearly, $\varphi(x) = 0$ for $x \leq s^*$. We now assume the induction hypothesis that $\varphi(x) \geq 0$ for $x \leq \bar{x}$ for some $\bar{x} \in [s^*, S^*]$. Then for any $x \in [\bar{x}, \min(\bar{x} + \eta, S^*)]$, $\varphi(x - D(p, \tilde{\varepsilon})) \geq 0$ following the induction hypothesis. Therefore we have,

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

where the last inequality follows from part (c). Thus, part (d) holds.

D Proof for Lemma 5

We prove by induction that for any $x_0 \leq x_1$, $\lambda \in [0,1]$,

$$\varphi(x_\lambda) \geq (1-\lambda)\varphi(x_0) + \lambda\varphi(x_1) - \max\{1-\lambda, \lambda\}k, \tag{24}$$

where $x_\lambda = (1-\lambda)x_0 + \lambda x_1$.

Since $\varphi(x) = 0$ for $x \leq s^*$, the inequality (24) trivially holds for $x_1 \leq s^*$. We assume, as the induction hypothesis, that (24) holds for any $x_0 \leq x_1 \leq \bar{x}$ for some $\bar{x}$. We next show that the inequality continues to hold for any $x_0 \leq x_1 \leq \bar{x} + \eta$.

Let $p^*(x)$ be an optimal solution for problem (16) with $s = s^*$. We distinguish between three different cases.

(1) $x_0 > s^*$. In this case, let $p_\lambda = (1-\lambda)p^*(x_0) + \lambda p^*(x_1)$. We have

$$\begin{aligned}
\varphi(x_\lambda) &\geq \mathcal{G}[P(x_\lambda, p_\lambda; \tilde{\varepsilon}) + \gamma\varphi(x_\lambda - D(p_\lambda, \tilde{\varepsilon}))] - Q(s^*) \\
&\geq \mathcal{G}[(1-\lambda)(P(x_0, p^*(x_0); \tilde{\varepsilon}) + \gamma\varphi(x_0 - D(p^*(x_0), \tilde{\varepsilon}))) \\
&\quad +\lambda(P(x_1, p^*(x_1); \tilde{\varepsilon}) + \gamma\varphi(x_1 - D(p^*(x_1), \tilde{\varepsilon}))) - \max\{1-\lambda, \lambda\}k] - Q(s^*) \\
&\geq (1-\lambda)\mathcal{G}[P(x_0, p^*(x_0); \tilde{\varepsilon}) + \gamma\varphi(x_0 - D(p^*(x_0), \tilde{\varepsilon}))] \\
&\quad +\lambda\mathcal{G}[P(x_1, p^*(x_1); \tilde{\varepsilon}) + \gamma\varphi(x_1 - D(p^*(x_1), \tilde{\varepsilon}))] - Q(s^*) - \max\{1-\lambda, \lambda\}k \\
&= (1-\lambda)\varphi(x_0) + \lambda\varphi(x_1) - \max\{1-\lambda, \lambda\}k,
\end{aligned}$$

where the first inequality holds since p_λ is a feasible price; the second inequality follows from the monotonicity of the operator $\mathcal{G}$, the concavity of the function $P(S^*, p, \tilde{\varepsilon})$ in p, the fact that $x_0 - D(p^*(x_0), \tilde{\varepsilon}) \leq \bar{x}$ and $x_1 - D(p^*(x_1), \tilde{\varepsilon}) \leq \bar{x}$; and the induction assumption, and the last inequality holds since the operator $\mathcal{G}$ preserves concavity.

(2) $x_0 \leq s^*$ and $x_\lambda \leq S^*$. In this case, we have from Lemma 4 that $\varphi(x_0) = 0$, $\varphi(x_\lambda) \geq 0$ and $\varphi(x_1) \leq k$. Thus, the inequality (24) holds.

(3) $x_0 \leq s^* \leq S^* \leq x_\lambda$. In this case, x_λ can be expressed as the convex combination of S^* and x_1, i.e., $x_\lambda = (1-\mu)S^* + \mu x_1$ for some $\mu \in [0,1]$. It is straightforward to verify that

$\mu \leq \lambda$. We now have that

$$\begin{aligned}
\varphi(x_\lambda) &\geq (1-\mu)\varphi(S^*) + \mu\varphi(x_1) - \max\{1-\mu, \mu\}k \\
&= (1-\lambda)\varphi(x_0) + \lambda\varphi(x_1) - \max\{1-\lambda, \lambda\}k \\
&\quad +(1-\mu)k + (\mu-\lambda)\varphi(x_1) \\
&\quad + \max\{1-\lambda, \lambda\}k - \max\{1-\mu, \mu\}k \\
&\geq (1-\lambda)\varphi(x_0) + \lambda\varphi(x_1) - \max\{1-\lambda, \lambda\}k \\
&\quad +(1-\mu)k + (\mu-\lambda)k \\
&\quad + \max\{1-\lambda, \lambda\}k - \max\{1-\mu, \mu\}k \\
&\geq (1-\lambda)\varphi(x_0) + \lambda\varphi(x_1) - \max\{1-\lambda, \lambda\}k,
\end{aligned}$$

where the first inequality follows from case (1), the second inequality from the fact that $\varphi(x) \leq k$ for any x and the last inequality from a simple algebraic manipulation.

Thus, the inequality (24) holds for $x_0 \leq x_1 \leq \bar{x} + \eta$ and the proof is complete.

E Proof for Theorem 3.2

Define for any x,

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

It is clear that $O(x) = \varphi(x)$ for $x \geq s^*$, while for $x \leq s^*$, $\varphi(x) = 0$ and $O(x) = Q(x) - Q(s^*) \leq 0$. Following Lemma 4, for any x we have

$$O(x) \leq O(S^*) = \varphi(S^*) = k \ .$$

We distinguish between three different cases to show the result.

(1) $x \leq s^*$. In this case, we have

$$\begin{aligned}
\varphi(x) = 0 &= -k + \max_{y: y \geq x} O(y) \\
&= \max_{y: y \geq x} -k\delta(y-x) + \max_{\underline{p} \leq p \leq \bar{p}} \mathcal{G}\left[P(y, p, \tilde{\varepsilon}) - Q(s^*) + \gamma\varphi(y - D(p, \tilde{\varepsilon}))\right] \ ,
\end{aligned}$$

with the maximum achieved at $y^* = S^*$.

(2) $x \in (s^*, S^*]$. In this case, $\varphi(x) \geq 0$, and for any $y > x$, $-k\delta(y-x) + O(y) \leq 0$. Therefore

$$\varphi(x) = O(x) = \max_{y: y \geq x} -k\delta(y-x) + \max_{\underline{p} \leq p \leq \bar{p}} \mathcal{G}\left[P(y, p, \tilde{\varepsilon}) - Q(s^*) + \gamma\varphi(y - D(p, \tilde{\varepsilon}))\right] \ ,$$

with the maximum achieved at $y^* = x$.

(3) $x > S^*$. In this case, for any $y > x$, x can be expressed as the convex combination of S^* and y, i.e., there exists $\lambda \in [0,1]$ such that $x = (1-\lambda)S^* + \lambda y$. Since φ is symmetric k-concave, we have that

$$\begin{aligned}\varphi(x) &\geq (1-\lambda)\varphi(S^*) + \lambda\varphi(y) - \max\{1-\lambda, \lambda\}k \\ &= \varphi(y) + (1-\lambda)(\varphi(S^*) - \varphi(y)) - \max\{1-\lambda, \lambda\}k \\ &\geq \varphi(y) - k \\ &= O(y) - k,\end{aligned}$$

where the first inequality follows from the symmetric k-concavity of φ and the last inequality follows from the fact that $\varphi(y) \leq \varphi(S^*) = k$ for any y. Thus, it is optimal not to place an order, or, $y^* = x$.

F Proof for Theorem 3.4

(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 $x \leq S^*$ (Lemma 4(b)(d)). The induction hypothesis is valid.

Let

$$y = y(x) = \begin{cases} x & x \geq s^* \\ S^* & x < s^* \end{cases} .$$

We have

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

Therefore

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

For any given inventory level $x > S^*$, since the inventory level will decrease to no more than S^* in finite number of periods, due to the demand being lower bounded by η, the point-wise convergence result still holds.

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

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

where y is the same as defined in part (a), and the first inequality follows from the definition of $y, p^*(y), G_t(x)$ 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. Again, using induction we can show that for any $x \leq \Delta$,

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

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

The proof is now complete.

G Additional results for the finite horizon models

Theorem G.1 *For the ambiguity and risk averse inventory and pricing model (8),*

(a) an (s, S, A, p) policy is optimal, and

(b) an (s, S) policy is optimal when pricing is not a decision variable.

The result is similar to the one corresponding to the risk averse inventory and pricing model analyzed in Chen et al. (2007). The proof is similar as well by observing that the minimum envelope of (symmetric) k-concave functions is still (symmetric) k-concave (see Proposition 7 in Appendix A) and thus is omitted.

We now further restrict the demand function $D_t(p_t, \tilde{\varepsilon}_t)$ and the ambiguity set Θ_t and show that the structure of the optimal inventory and pricing policy may be reduced to an (s, S, p) policy.

Theorem G.2 *Consider the risk neutral ambiguity averse inventory and pricing formulation, i.e., Equation (8) with $\rho_t \to \infty$. If*

- *the demand function is additive, i.e., parameter α_t in the demand function*

$$D_t(p_t, \tilde{\varepsilon}_t) = \tilde{\beta}_t - \alpha_t p_t$$

  is a constant; and

- *every $\tilde{\beta}_t$ in set Θ_t has the same expectation,*

then an (s, S, p) policy is optimal.

The result and proof are similar to Chen and Simchi-Levi (2004a), Theorem 3.1. For the completeness of the paper, here we provide a key step for the proof, which basically shows that a higher order-up-to level y_t leads to a higher expected end-period inventory level.

Proof. Note that (8) under the conditions specified in the theorem can be written as follows.

$$J_t(x) = \max_{y:y\geq x} -k\delta(y-x) + \max_{p:\underline{p}_t \leq p \leq \bar{p}_t} I_t(y,p) \ ,$$

where

$$I_t(y,p) = R(p) + F_t(y + \alpha_t p),$$

$$R(p) = (p - c_t)(\beta_t - \alpha_t p) \ ,$$

with β_t being the expectation of $\tilde{\beta}_t$, and

$$F_t(y) = \min_{f_{\tilde{\beta}_t} \in \Theta_t} E_{f_{\tilde{\beta}_t}} \left[-\hat{h}(y - \tilde{\beta}_t) + \gamma J_{t-1}(y - \tilde{\beta}_t) \right] \ .$$

Define a new function $K(y,z)$ as follows.

$$K(y,z) = R((z-y)/\alpha_t) + F_t(z).$$

Then, we have that

$$\max_{p:\underline{p}_t \leq p \leq \bar{p}_t} I_t(y,p) = \max_{z:\underline{p}_t \leq (z-y)/\alpha_t \leq \bar{p}_t} K(y,z).$$

Since $R(\cdot)$ is concave and $K(y,z)$ is supermodular, the above optimization problem has an optimal solution $z^*(y)$, which is non-deceasing in y. This implies that the higher the order-up-to level y in period t, the higher the expected inventory level at the end of period t. This result is parallel to Lemma 2 in Chen and Simchi-Levi (2004a). The remaining proof follows from the same steps as the one for Theorem 3.1 in Chen and Simchi-Levi (2004a) and is omitted. ■

From Theorem G.1, we know that an (s, S, A, p) policy is optimal for the finite horizon model (8), or,

$$G_t(x) = \max_{y \geq x, p \in \mathcal{P}} -k\delta(y-x) + \mathcal{G}_{\Theta}^{R_t}\left[P(y, p; \tilde{\varepsilon}) + \gamma G_{t-1}(y - D(p, \tilde{\varepsilon}))\right] .$$

Next we present a technical result which illustrates that the parameters s_t and S_t are uniformly bounded for $t = 1, 2, \ldots$. For this purpose, let the lower bound $\underline{s}$ be a constant such that $\underline{s} \leq y^*$ and

$$\hat{h}(\underline{s}) = \hat{h}(y^*) + k,$$

where y^* is a minimizer of $\hat{h}$.

Also define the upper bound to be $\bar{S} = S^0 + \Xi$, in which $S^0 \geq y^*$ and $\hat{h}(S^0) = \hat{h}(y^*) + k$, and Ξ is the demand upper bound in Assumption 2.1. Indeed, the demand upper bound assumption is only used here to derive an upper bound for S_t. It is worth noticing that in the risk neutral case, such an assumption is not needed (see Chen and Simchi-Levi (2004b)).

The main result in this subsection is the following theorem.

Theorem G.3 *$\underline{s}$ and $\bar{S}$ are the lower and upper bound for s_t and S_t. That is,*

$$\underline{s} \leq s_t \leq S_t \leq \bar{S} .$$

In order to prove Theorem G.3, we first make the following observations.

Lemma 8 *(a) $G_t(x) \geq G_t(x') - k$ for $x \leq x'$.*

(b) $G_t(x) \leq G_t(x')$ for any $x \leq x' \leq y^*$.

(c) $g_t(y, p) \leq g_t(y', p)$ for any $y \leq y' \leq y^*$ and $p \in \mathcal{P}$, where

$$g_t(y, p) = \mathcal{G}_{\Theta}^{R_t}\left[P(y, p; \tilde{\varepsilon}) + \gamma G_{t-1}(y - D(p, \tilde{\varepsilon}))\right].$$

Proof. For two inventory levels $x \leq x'$, one can always first raise the inventory level from x to x' by paying a fixed cost k and then follow the same optimal strategy for the inventory level x'. Therefore (a) holds.

We now prove parts (b) and (c) by induction. First, we have that part (b) holds for $t = 0$ since $G_0(x) = 0$. Now assume that (b) holds for $G_{t-1}(x)$. Then for any $y \leq y' \leq y^*$ and $p \in \mathcal{P}$,

$$\begin{aligned} P(y, p; \tilde{\varepsilon}) &= (p-c)D(p, \tilde{\varepsilon}) - \hat{h}(y' - D(p, \tilde{\varepsilon})) + \hat{h}(y' - D(p, \tilde{\varepsilon})) - \hat{h}(y - D(p, \tilde{\varepsilon})) \\ &\leq P(y', p; \tilde{\varepsilon}), \end{aligned}$$

where the inequality holds since $\hat{h}(x)$ is convex and $y \le y' \le y^*$. Thus, for any $y \le y' \le y^*$ and $p \in \mathcal{P}$,

$$\begin{aligned} g_t(y,p) &= \mathcal{G}_{\Theta}^{R_t}\left[P(y,p;\tilde{\varepsilon}) + \gamma G_{t-1}(y - D(p,\tilde{\varepsilon}))\right] \\ &\le \mathcal{G}_{\Theta}^{R_t}\left[P(y',p;\tilde{\varepsilon}) + \gamma G_{t-1}(y' - D(p,\tilde{\varepsilon}))\right] \\ &= g_t(y',p), \end{aligned}$$

where the inequality follows from the induction assumption. Thus, part (c) holds. Finally, part (b) follows immediately from part (c), and the fact that $G_t(x) = \max_p g_t(x,p)$. ■

Now we are ready to prove Theorem G.3.

Proof. First, note that for any $p \in \mathcal{P}$ and $y \le \underline{s}$, we have that

$$\begin{aligned} P(y,p;\tilde{\varepsilon}) &= (p-c)D(p,\tilde{\varepsilon}) - \hat{h}(y^* - D(p,\tilde{\varepsilon})) + \hat{h}(y^* - D(p,\tilde{\varepsilon})) - \hat{h}(y - D(p,\tilde{\varepsilon})) \\ &\le P(y^*,p;\tilde{\varepsilon}) + \hat{h}(y^*) - \hat{h}(y) \\ &\le P(y^*,p;\tilde{\varepsilon}) - k, \end{aligned}$$

where the first inequality holds since $\hat{h}$ is convex and thus has increasing differences, and the second inequality follows from the definition of $\underline{s}$. In addition, from Lemma 8 (b), we have $G_{t-1}(y - D(p,\tilde{\varepsilon})) \le G_{t-1}(y^* - D(p,\tilde{\varepsilon}))$ for $y \le \underline{s} \le y^*$. Therefore,

$$\begin{aligned} g_t(y,p) &= \mathcal{G}_{\Theta}^{R_t}\left[P(y,p;\tilde{\varepsilon}) + \gamma G_{t-1}(y - D(p,\tilde{\varepsilon}))\right] \\ &\le \mathcal{G}_{\Theta}^{R_t}\left[P(y^*,p;\tilde{\varepsilon}) + \gamma G_{t-1}(y^* - D(p,\tilde{\varepsilon}))\right] - k \\ &= g_t(y^*,p) - k. \end{aligned}$$

This implies that it is optimal to place an order for an inventory level no more than $\underline{s}$.

We now show that $S_t \le \bar{S}$. Note that for any $p \in \mathcal{P}$ and $y \ge \bar{S}$, we have that

$$\begin{aligned} P(y,p;\tilde{\varepsilon}) &= (p-c)D(p,\tilde{\varepsilon}) - \hat{h}(y^* + \Xi - D(p,\tilde{\varepsilon})) + \hat{h}(y^* + \Xi - D(p,\tilde{\varepsilon})) - \hat{h}(y - D(p,\tilde{\varepsilon})) \\ &\le P(y^* + \Xi, p;\tilde{\varepsilon}) + \hat{h}(y^*) - \hat{h}(y - \Xi) \\ &\le P(y^* + \Xi, p;\tilde{\varepsilon}) - k, \end{aligned}$$

where again the first inequality holds since $\hat{h}$ is convex and the second inequality follows from the definition of $\bar{S}$. Thus, for any $y \ge \bar{S}$,

$$\begin{aligned} g_t(y,p) &= \mathcal{G}_{\Theta}^{R_t}\left[P(y,p;\tilde{\varepsilon}) + \gamma G_{t-1}(y - D(p,\tilde{\varepsilon}))\right] \\ &\le \mathcal{G}_{\Theta}^{R_t}\left[P(y^* + \Xi,p;\tilde{\varepsilon}) - k + \gamma G_{t-1}(y - D(p,\tilde{\varepsilon}))\right] \\ &\le \mathcal{G}_{\Theta}^{R_t}\left[P(y^* + \Xi,p;\tilde{\varepsilon}) - k + \gamma G_{t-1}(y^* + \Xi - D(p,\tilde{\varepsilon})) + \gamma k\right] \\ &\le g_t(y^* + \Xi, p), \end{aligned}$$

where the second inequality follows from Lemma 8 (a). Thus, $S_t \le \bar{S}$. ■

H Concluding remarks and additional discussions

In a finite planning horizon, an (s, S, A, p) policy is optimal for our problem. Interestingly, in the stationary infinite horizon case, the optimal policy structure simplifies to an (s, S, p) policy for both the discounted and the long run average optimization criteria. If we restrict to the ambiguity averse models, under certain conditions, an (s, S, p) policy is also optimal to the finite horizon case.

Our proof for the infinite horizon case is much more complicated than the corresponding one for the risk neutral model in Chen and Simchi-Levi (2004b), even though they bear some similarities. 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 then illustrate that the value function of the best (s, S) inventory policy gives a solution to the Bellman equation. 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. Another point that differs our proof from Chen and Simchi-Levi (2004b) lies in the derivation of the upper bounds for the parameters s_t and S_t in the optimal (s, S, A, p) policy for the finite horizon model. Here we assume that there is a uniform upper bound on realized demand, which was not assumed in the proof of the risk neutral model.

It is worth mentioning that for the discounted infinite horizon inventory and pricing problem with risk neutrality and no ambiguity, Huh and Janakiraman (2004) 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.

Finally, another 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 exist even in an ambiguity averse newsvendor setting.

References

Chen, X. (2003). *Coordinating inventory control and pricing strategies with random demand and fixed ordering cost.* (Ph.D. thesis, Massachusetts Institute of Technology)

Chen, X., Sim, M., Simchi-levi, D., & Sun, P. (2007). Risk aversion in inventory management. *Operations Research*, *55*(5), 828-842.

Chen, X., & Simchi-Levi, D. (2004a). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research*, *52*(6), 887-896.

Chen, X., & Simchi-Levi, D. (2004b). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Mathematics of Operations Research*, *29*(3), 698-723.

Huh, W. T., & Janakiraman, G. (2004). Alternative approaches to important results in inventory and pricing control. *Operations Research*. (To appear)

Pratt, J. (1964). Risk aversion in the small and in the large. *Econometrica*, *32*(1-2), 122-136.

Rockafellar, R. T. (1970). *Convex analysis.* Princeton University Press, Princeton.

Scarf, H. (1958). A min-max solution to an inventory problem. In K. Arrow, S. Karlin, & H. Scarf (Eds.), *Studies in mathematical theory of inventory and production* (p. 201-209). Stanford University Press, Stanford, CA.

Scarf, H. (1960). The optimality of (s, s) policies for the dynamic inventory problem. In *Proceedings of the 1st stanford symposium on mathematical methods in the social sciences.* Stanford University Press, Stanford, CA.

Simchi-Levi, D., Chen, X., & J., B. (2004). *The logic of logistics: theory, algorithms, and applications for logistics and supply chain management, second edition.* Springer-Verlag, New York.