Stochastic Stackelberg equilibria with applications to time-dependent newsvendor models

Bernt Øksendal[a,1,*], Leif Sandal[b,2], Jan Ubøe[b]

[a] *Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, 0316 Oslo, Norway*
[b] *Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway*

Abstract

In this paper, we prove a maximum principle for general stochastic differential Stackelberg games, and apply the theory to continuous time newsvendor problems. In the newsvendor problem, a manufacturer sells goods to a retailer, and the objective of both parties is to maximize expected profits under a random demand rate. Our demand rate is an Itô–Lévy process, and to increase realism information is delayed, e.g., due to production time. A special feature of our time-continuous model is that it allows for a price-dependent demand, thereby opening for strategies where pricing is used to manipulate the demand.
Keywords: stochastic differential games, delayed information, Itô-Lévy processes, Stackelberg equilibria, newsvendor models, optimal control of forward-backward stochastic differential equations

*Corresponding author

Email addresses: oksendal@math.uio.no (Bernt Øksendal), leif.sandal@nhh.no (Leif Sandal), jan.uboe@nhh.no (Jan Ubøe)

[1]The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no [228087]

[2]The research leading to these results has received funding from NFR project 196433

Main variables:

$w =$ wholesale price per unit (chosen by the manufacturer)

$q =$ order quantity (rate chosen by the retailer)

$R =$ retail price per unit (chosen by the retailer)

$D =$ demand (random rate)

$M =$ production cost per unit (fixed)

$S =$ salvage price per unit (fixed)

1. Introduction

The one-period newsvendor model is a widely studied object that has attracted increasing interest in the last two decades. The basic setting is that a retailer wants to order a quantity q from a manufacturer. Demand D is a random variable, and the retailer wishes to select an order quantity q maximizing his expected profit. When the distribution of D is known, this problem is easily solved. The basic problem is very simple, but appears to have a never-ending number of variations. There is now a very large literature on such problems, and for further reading we refer to the survey papers by Cachón (2003) and Qin et al. (2011) and the numerous references therein.

The (discrete) multiperiod newsvendor problem has been studied in detail by many authors, including Matsuyama (2004), Berling (2006), Bensoussan et al. (2007, 2009), Wang et al., (2010), just to quote some of the more recent contributions. Two papers whose approach is not unlike that used in our paper are Kogan (2003) and Kogan and Lou (2003), where the authors consider continuous time-scheduling problems.

In many cases, demand is not known and the parties gain information through a sequence of observations. There is a huge literature on cases with partial information, e.g., Scarf (1958), Gallego & Moon (1993), Bensoussan et al. (2007), Perakis & Roels (2008), Wang et al. (2010), just to mention a few. When a sufficiently large number of observations have been made, the distribution of demand is fully revealed and can be used to optimize order quantities. This approach only works if the distribution of D is static, and leads to false conclusions if demand changes systematically over time. In this paper we will assume that the demand rate is a stochastic process D_t and we seek optimal decision rules for that case.

In our paper, a retailer and a manufacturer write contracts for a specific delivery rate following a decision process in which the manufacturer is the leader who initially decides the wholesale price. Based on that wholesale price, the retailer decides on the delivery rate and the retail price. We assume a Stackelberg framework, and hence ignore cases where the retailer can negotiate the wholesale price. The contract is written at time $t-\delta$, and goods are received at time t. It is essential to assume that information is delayed. If there is no delay, the demand rate is known, and the retailer's order rate is made equal to the demand rate. Information is delayed by a time δ. One justification for this is that production takes time, and orders cannot be placed and effectuated instantly. It is natural to think about δ as a production lead time.

The single period newsvendor problem with price dependent demand is classical, see Whitin (1955). Mills (1959) refined the construction considering the case where demand uncertainty is added to the price-demand curve, while Karlin and Carr (1962) considered the case where demand uncertainty is multiplied with the price-demand curve. For a nice review of the problem with extensions see Petruzzi and Dada (1999). Stackelberg games for single period newsvendor problems with fixed retail price have been studied extensively by Lariviere and Porteus (2001), providing quite general conditions under which unique equilibria can be found.

Multiperiod newsvendor problems with delayed information have been discussed in several papers, but none of these papers appears to make the theory operational. Bensoussan et al. (2009) use a time-discrete approach and generalize several information delay models. However, these are all under the assumption of independence of the delay process from inventory, demand, and the ordering process. They assert that removing this assumption would give rise to interesting as well as challenging research problems, and that a study of computation of the optimal base stock levels and their behavior with respect to problem parameters would be of interest. Computational issues are not explored in their paper, and they only consider decision problems for inventory managers, disregarding any game theoretical issues.

Calzolari et al. (2011) discuss filtering of stochastic systems with fixed delay, indicating that problems with delay lead to nontrivial numerical difficulties even when the driving process is Brownian motion. In our paper, solutions to general delayed newsvendor equilibria are formulated in terms of coupled systems of stochastic differential equations. Our approach may hence be useful also in the general case where closed form solutions cannot be obtained.

Stochastic differential games have been studied extensively in the literature. However, most of the works in this area have been based on dynamic programming and the associated Hamilon-Jacobi-Bellman-Isaacs type of equations for systems driven by Brownian motion only. More recently, papers on stochastic differential games based on the maximum principle (including jump diffusions) have appeared. See, e.g., Øksendal and Sulem (2012) and the references therein. This is the approach used in our paper, and as far as we know, the application to the newsvendor model is new. The advantage with the maximum approach is two-fold:

- We can handle non-Markovian state equations and non-Markovian payoffs.

- We can deal with games with partial and asymmetric information.

Figure 1 shows a sample path of an Ornstein–Uhlenbeck process that is mean reverting around a level $\mu = 100$. Even though the long-time average is 100, orders based on this average are clearly suboptimal. At, e.g., $t = 30$, we observe a demand rate $D_{30} = 157$. When the mean reversion rate is as slow as in Figure 1, the information $D_{30} = 157$ increases the odds that the demand rate is more than 100 at time $t = 37$. If the delay $\delta = 7$ (days), the retailer should hence try to exploit this extra information to improve performance.

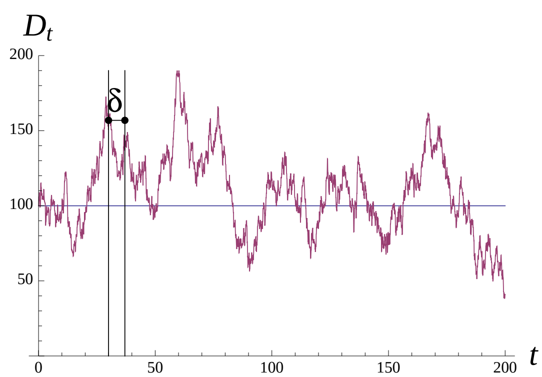


Figure 1: An Ornstein–Uhlenbeck process with delayed information

Based on the information available at time $t - \delta$, the manufacturer should offer the retailer a price per unit w_t for items delivered at time t. Given the wholesale price w_t and all available information, the retailer should decide on an order rate q_t and a retail price R_t. The retail price can in principle lead to changes in demand, and in general the demand rate D_t is, hence, a function of R_t. However, such cases are hard to solve in terms of explicit expressions. We will also look at the simplified case where R is exogenously given and fixed. To carry out our construction, we will need to assume that items cannot be stored. That is of course a strong limitation, but applies to important cases like electricity markets and markets for fresh foods.

Assuming that both parties have full information about demand rate at time $t-\delta$, and that the manufacturer knows how much the retailer will order at any given unit price w, we are left with a Stackelberg game where the manufacturer is the leader and the retailer is the follower. To our knowledge, stochastic differential games of this sort have not been discussed in the literature previously. Before we can discuss game equilibria for the newsvendor problem, we must formulate and prove a maximum principle for general stochastic differential Stackelberg games.

In the case where R is exogenously given and fixed, it seems reasonable to conjecture that our optimization problem could be reduced to solving a family of static newsvendor problems pointwise in t. Theorem 3.2.2 confirms that this approach provides the correct solution to the problem. Note, however, that our general framework is non-Markovian, and that solutions may depend on path properties of the demand.

The paper is organized as follows. In Section 2, we set up a framework where we discuss general stochastic differential Stackelberg games. In Section 3, we use the machinery in Section 2 to consider a continuous-time newsvendor problem. In Section 4, we consider the special case where the demand rate is given by an Ornstein–Uhlenbeck process and provide explicit solutions for the unique equilibria that occur in that case. Examples with R-dependent demand are considered in Section 5. Finally, in Section 6 we offer some concluding remarks.

2. General stochastic differential Stackelberg games

In this section, we will consider general stochastic differential Stackelberg games. In our framework, the state of the system is given by a stochastic process X_t. The game has two players. Player 1 (leader, denoted by L) can at time t choose a control $u_L(t)$ while player 2 (follower, denoted by F) can choose a control $u_F(t)$. The controls determine how X_t evolves in time. The performance for player i is assumed to be of the form

$$J_i(u_L, u_F) = \mathrm{E}\left[\int_\delta^T f_i(t, X_t, u_L(t), u_F(t), \omega)dt + g_i(X_T, \omega)\right] \qquad i = L, F \tag{1}$$

where $f_i(t,x,w,v,\omega) : [0,T]\times\mathbb{R}\times\mathbb{R}^l\times\mathbb{R}^m\times\Omega\to\mathbb{R}$ is a given $\mathcal{F}_t$-adapted process and $g_i(x,\omega):\mathbb{R}\times\Omega\to\mathbb{R}$ are given $\mathcal{F}_T$-measurable random variables for each $x,w,v; i=L,F$. We will assume that f_i are C^1 in v,w,x and that g_i are C^1 in x, $i=L,F$.

In our Stackelberg game, player 1 is the leader, and player 2 the follower. Hence when u_L is revealed to the follower, the follower will choose u_F to maximize $J_F(u_L, u_F)$. The leader knows that the follower will act in this rational way.

Suppose that for any given control u_L there exists a map Φ (a "maximizer" map) that selects u_F that maximizes $J_F(u_L, u_F)$. The leader will hence choose $u_L = u_L^*$ such that $u_L \mapsto J_L(u_L, \Phi(u_L))$ is maximal for $u_L = u_L^*$. In order to solve problems of this type we need to specify how the state of the system evolves in time. We will assume that the state of the system is given by a controlled jump diffusion of the form:

$$\begin{aligned} dX_t &= \mu(t, X_t, u(t), \omega)dt + \sigma(t, X_t, u(t), \omega)dB_t \\ &\quad + \int_{\mathbb{R}} \gamma(t, X_{t^-}, u(t), \xi, \omega)\tilde{N}(dt, d\xi) \\ X(0) &= x \in \mathbb{R} \end{aligned} \tag{2}$$

where the coefficients $\mu(t, x, u, \omega) : [0, T] \times \mathbb{R} \times \mathbb{U} \times \Omega \to \mathbb{R}$, $\sigma(t, x, u, \omega) : [0, T] \times \mathbb{R} \times \mathbb{U} \times \Omega \to \mathbb{R} \times \mathbb{R}^n$, $\gamma(t, x, u, \xi, \omega) : [0, T] \times \mathbb{R} \times \mathbb{U} \times \mathbb{R}_0 \times \Omega \to \mathbb{R}$ are given continuous functions assumed to be continuously differentiable with respect to x and u, and $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$. Here $B_t = B(t, \omega);\ (t, \omega) \in [0, \infty) \times \Omega$ is a Brownian motion in $\mathbb{R}^n$ and $\tilde{N}(dt, d\xi) = \tilde{N}(dt, d\xi, \omega)$ is an independent compensated Poisson random measure on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$. See Øksendal and Sulem (2007) for more information about controlled jump diffusions. The set $\mathbb{U} = \mathbb{U}_L \times \mathbb{U}_F$ is a given set of admissible control values $u(t, \omega)$. We assume that the control $u = u(t, \omega)$ consists of two components, $u = (u_L, u_F)$, where the leader controls $u_L \in \mathbb{R}^l$ and the follower controls $u_F \in \mathbb{R}^m$. We also assume that the information flow available to the players is given by the filtration $\{\mathcal{E}_t\}_{t \in [0,T]}$, where

$$\mathcal{E}_t \subseteq \mathcal{F}_t \qquad \text{for all } t \in [0, T]. \tag{3}$$

For example, the case much studied in this paper is when

$$\mathcal{E}_t = \mathcal{F}_{t-\delta} \qquad \text{for all } t \in [\delta, T]. \tag{4}$$

for some fixed information delay $\delta > 0$. We assume that $u_L(t)$ and $u_F(t)$ are $\mathcal{E}_t$-predictable, and assume there is given a family $\mathcal{A}_{\mathcal{E}} = \mathcal{A}_{L,\mathcal{E}} \times \mathcal{A}_{F,\mathcal{E}}$ of admissible controls contained in the set of $\mathcal{E}_t$-predictable processes.

We now consider the following game theoretic situation:

Suppose the leader decides her control process $u_L \in \mathcal{A}_{L,\mathcal{E}}$. At any time t the value is immediately known to the follower. Therefore he chooses $u_F = u_F^* \in \mathcal{A}_{F,\mathcal{E}}$ such that

$$u_F \mapsto J_F(u_L, u_F) \text{ is maximal for } u_F = u_F^*. \tag{5}$$

Assume that there exists a measurable map $\Phi : \mathcal{A}_{L,\mathcal{E}} \to \mathcal{A}_{F,\mathcal{E}}$ such that

$$u_F \mapsto J_F(u_L, u_F) \text{ is maximal for } u_F = u_F^* = \Phi(u_L) \tag{6}$$

The leader knows that the follower will act in this rational way. Therefore the leader will choose $u_L = u_L^* \in \mathcal{A}_{L,\mathcal{E}}$ such that

$$u_L \mapsto J_L(u_L, \Phi(u_L)) \text{ is maximal for } u_L = u_L^*. \tag{7}$$

The control $u^* := (u_L^*, \Phi(u_L^*)) \in \mathcal{A}_{L,\mathcal{E}} \times \mathcal{A}_{F,\mathcal{E}}$ is called a *Stackelberg equilibrium* for the game defined by (1)-(2). In the newsvendor problem studied in this paper, the leader is the manufacturer who decides the wholesale price $u_L = w$ for the retailer, who is the follower, and who decides the order rate $u_F^{(1)} = q$ and the retailer price $u_F^{(2)} = R$. Thus $u_F = (q, R)$. We may summarize (5) and (7) as follows:

$$\max_{u_F \in A_{F,\mathcal{E}}} J_F(u_L, u_F) = J_F(u_L, \Phi(u_L)) \tag{8}$$

and

$$\max_{u_L \in \mathcal{A}_{L,\mathcal{E}}} J_L(u_L, \Phi(u_L)) = J_L(u_L^*, \Phi(u_L^*)) \tag{9}$$

We see that (8) and (9) constitute two consecutive stochastic control problems with partial information, and hence we can, under some conditions, use the maximum principle for such problems as presented in Øksendal and Sulem (2012) (see also, e.g., Framstad et al. (2004) and Baghery and Øksendal (2007)) to find a maximum principle for Stackelberg equilibria. To this end, we define the Hamiltonian $H_F(t, x, u, a_F, b_F, c_F(\cdot), \omega) : [0,T] \times \mathbb{R} \times \mathbb{U} \times \mathbb{R} \times \mathbb{R}^{n+1} \times \mathcal{R} \times \Omega \to \mathbb{R}$ by

$$\begin{aligned} H_F(t, x, u, a_F, b_F, c_F(\cdot), \omega) = f_F(t, x, u, \omega) + \mu(t, x, u, \omega) a_F + \sigma(t, x, u, \omega) b_F \\ + \int_{\mathbb{R}} \gamma(t, x, u, \xi, \omega) c_F(\xi) \nu(d\xi); \end{aligned} \tag{10}$$

where $\mathcal{R}$ is the set of functions $c(\cdot) : \mathbb{R}_0 \to \mathbb{R}$ such that (10) converges, ν is a Lévy measure. For simplicity of notation the explicit dependence on $\omega \in \Omega$ is suppressed in the following. The adjoint equation for H_F in the unknown adjoint processes $a_F(t), b_F(t)$, and $c_F(t,\xi)$ is the following backward stochastic differential equation (BSDE):

$$da_F(t) = -\frac{\partial H_F}{\partial x}(t, X(t), u(t), a_F(t), b_F(t), c_F(t,\cdot))dt + b_F(t)dB_t + \int_{\mathbb{R}} c_F(t,\xi)\tilde{N}(dt,d\xi); \qquad 0 \leq t \leq T \tag{11}$$

$$a_F(T) = g_F{}'(X(T)) \tag{12}$$

Here $X(t) = X^u(t)$ is the solution to (2) corresponding to the control $u \in \mathcal{A}_{\mathcal{E}}$. Next, assume that there exists a function $\phi : [0,T] \times \mathbb{U}_L \times \Omega \to \mathbb{U}_F$ such that

$$\Phi(u_L)(t) = \phi(t, u_l(t)) \qquad \text{i.e} \qquad \Phi(u_L) = \phi(\cdot, u_L(\cdot)) \tag{13}$$

Define the Hamiltonian $H_L^{\phi}(t, x, u_L, a_L, b_L, c_L(\cdot)) : [0,T] \times \mathbb{R} \times \mathbb{U}_L \times \mathbb{R} \times \mathbb{R}^{n+1} \times \mathcal{R} \to \mathbb{R}$ by

$$H_L^{\phi}(t, x, u_L, a_L, b_L, c_L(\cdot)) = f_L(t, x, u_L, \Phi(u_L)) + \mu(t, x, u_L, \Phi(u_L))a_L + \sigma(t, x, u_L, \Phi(u_L))b_L + \int_{\mathbb{R}} \gamma(t, x, u_L, \Phi(u_L), \xi)c_L(\xi)\nu(d\xi) \tag{14}$$

The adjoint equation (for H_L^{ϕ}) in the unknown processes $a_L(t), b_L(t), c_L(t,\xi)$ is the following BSDE:

$$da_L(t) = -\frac{\partial H_L^{\phi}}{\partial x}(t, X(t), u_L(t), \phi(t, u_L(t)), a_L(t), b_L(t), c_L(t,\cdot))dt + b_L(t)dB_t + \int_{\mathbb{R}} c_L(t,\xi)\tilde{N}(dt,d\xi); \qquad 0 \leq t \leq T \tag{15}$$

$$a_L(T) = g_L{}'(X(T)) \tag{16}$$

Here $X(t) = X^{u_L, \Phi(u_L)}(t)$ is the solution to (2) corresponding to the control $u(t) := (u_L(t), \phi(t, u_L(t)))$; $t \in [0,T]$, assuming that this is admissible.

We make the following assumptions:

(A1) For all $u_i \in \mathcal{A}_{i,\mathcal{E}}$ and all bounded $\beta_i \in \mathcal{A}_{i,\mathcal{E}}$ there exists $\epsilon > 0$ such that

$$u_i + s\beta_i \in \mathcal{A}_{i,\mathcal{E}} \text{ for all } s \in (-\epsilon, \epsilon);\ i = L, F.$$

(A2) For all $t_0 \in [0,T]$ and all bounded $\mathcal{E}_{t_0}$-measurable random variables α_i, the control process $\beta_i(t)$ defined by

$$\beta_i(t) = \begin{cases} \alpha_i & \text{if } t \in [t_0, T] \\ 0 & \text{otherwise} \end{cases} ; \; t \in [0,T]$$

belongs to $\mathcal{A}_{i,\mathcal{E}}$; $i = L, F$.

(A3) For all $u_i, \beta_i \in \mathcal{A}_{i,\mathcal{E}}$ with β_i bounded, the derivative processes

$$\xi_L(t) = \frac{d}{ds}\left(X^{u_L+s\beta_L,u_F}(t)\right)\Big|_{s=0}$$

$$\xi_F(t) = \frac{d}{ds}\left(X^{u_L,u_F+s\beta_F}(t)\right)\Big|_{s=0}$$

exist and belong to $L^2(\lambda \times P)$, where λ denotes Lebesgue measure on $[0,T]$.

We can now formulate our maximum principle for Stackelberg equilibria:

Theorem 2.1 (Maximum principle)

Assume that (13) and (A1)–(A3) hold. Put $u = (u_L, u_F) = (u_L, \Phi(u_L))$ where $\Phi : \mathbb{U}_L \to \mathbb{U}_F$, and let $X(t), (a_i, b_i, c_i)$ be the corresponding solutions of (2), (11)–(12) (for $i = F$) and (15)–(16) (for $i = L$), respectively.

Suppose that for all bounded $\beta_i \in \mathcal{A}_{i,\mathcal{E}}$, $i = L, F$ we have

$$\begin{aligned} \mathrm{E}\Big[\int_0^T \Big\{ &(a_i(t))^2 \Big(\Big(\frac{\partial \sigma}{\partial x}(t)\xi_i(i) + \frac{\partial \sigma}{\partial u_i}(t)\beta_i(t)\Big)^2 \\ &+ \int_{\mathbb{R}} \Big(\frac{\partial \gamma}{\partial x}(t,\zeta)\xi_i(t) + \frac{\partial \gamma}{\partial u_i}(t,\zeta)\beta_i(t)\Big)^2 \nu(d\zeta)\Big) \\ &+ \xi_i^2(t)\Big((b_i(t))^2 + \int_{\mathbb{R}} (c_i(t,\zeta))^2 \nu(d\zeta)\Big)\Big\} dt\Big] < \infty \end{aligned} \tag{17}$$

Then the following, (I) and (II), are equivalent.

(I)

$$\frac{d}{ds}\left(J_F(u_L, \Phi(u_L) + s\beta_F)\right)\Big|_{s=0} = \frac{d}{ds}\left(J_L(u_L + s\beta_L, \Phi(u_L + s\beta_L))\right)\Big|_{s=0} = 0 \tag{18}$$

for all bounded $\beta_L \in \mathcal{A}_{L,\mathcal{E}}, \beta_F \in \mathcal{A}_{F,\mathcal{E}}$.

(II)

$$\mathrm{E}\left[\frac{\partial}{\partial v_F} H_F(t, X(t), u_L(t), v_F, a_F(t), b_F(t), c_F(t,\cdot)) \Big| \mathcal{E}_t\right]_{v_F = \Phi(u_L)} = 0 \tag{19}$$

for $j = 1, 2$ and

$$\mathrm{E}\left[\frac{\partial}{\partial v_L} H_L^{\phi}(t, X(t), v_L, a_L(t), b_L(t), c_L(t, \cdot)) \middle| \mathcal{E}_t\right]_{v_L = u_L(t)} = 0 \tag{20}$$

Proof

This follows by first applying the maximum principle for optimal control with respect to $u_F \in \mathcal{A}_{F,\mathcal{E}}$ of the state process $X^{u_L,u_F}(t)$ for fixed $u_L \in \mathcal{A}_{L,\mathcal{E}}$, as presented in Øksendal and Sulem (2012). See also Framstad et al. (2004), Baghery and Øksendal (2007), Øksendal and Sulem (2007). Next we apply the same maximum principle with respect to $u_L \in \mathcal{A}_{L,\mathcal{E}}$ of the state process $X^{u_L,\Phi(u_L)}(t)$, for the given function Φ. We omit the details.

□

Corollary 2.2

Suppose $(u_L, \Phi(u_L))$ is a Stackelberg equilibrium for the game (1)-(2) and that (13), (A1)-(A3), and (17) are satisfied. Then the first order conditions (19)–(20) hold.

3. A continuous time newsvendor problem

In this section, we will formulate a continuous time newsvendor problem and use the results in Section 2 to describe a set of explicit equations that we need to solve to find Stackelberg equilibria. We will assume that the demand rate for a good is given by a (possibly controlled) stochastic process D_t. A retailer is at time $t-\delta$ offered a unit price w_t for items to be delivered at time t. Here $\delta > 0$ is the delay time. At time $t-\delta$, the retailer chooses an order rate q_t. The retailer also decides a retail price R_t. We assume that items can be salvaged at a unit price $S \geq 0$, and that items cannot be stored, i.e., they must be sold instantly or salvaged.

Remarks

The delay δ can be interpreted as a production lead time, and it is natural to assume that w_t and q_t should both be settled at time $t-\delta$. In general the retail price R_t can be settled at a later stage. To simplify notation we assume that R_t, too, is settled at time $t-\delta$. The assumption that items cannot be stored is, of course, quite restrictive. Many important cases lead to assumptions of this kind; we mention in particular the electricity market and markets for fresh foods.

Assuming that sale will take part in the time period $\delta \leq t \leq T$, the retailer will get an expected profit

$$J_F(w,q,R) = \mathrm{E}\left[\int_\delta^T (R_t - S)\min[D_t, q_t] - (w_t - S)q_t dt\right] \tag{21}$$

When the manufacturer has a fixed production cost per unit M, the manufacturer will get an expected profit

$$J_L(w,q,R) = \mathrm{E}\left[\int_\delta^T (w_t - M)q_t dt\right] \tag{22}$$

Technical remarks

To solve these problems mathematically, it is convenient to apply an equivalent mathematical formulation: At time t the retailer orders the quantity t for *immediate* delivery, but the information at that time is the delayed information $\mathcal{F}_{t-\delta}$ about the demand δ units of time. Similarly, when the manufacturer delivers the ordered quantity q_t at time t, the unit price w_t is based on $\mathcal{F}_{t-\delta}$. From a practical point of view this formulation is entirely different, but leads to the same optimization problem.

3.1. Formalized information

We will assume that our demand rate is given by a (possibly controlled) process of the form

$$dD_t = \mu(t, D_t, R_t, \omega)dt + \sigma(t, D_t, R_t, \omega)dB_t + \int_{\mathbb{R}} \gamma(t, D_{t^-}, R_t, \xi, \omega)\tilde{N}(dt, d\xi); \qquad t \in [0,T] \tag{23}$$
$$D_0 = d_0 \in \mathbb{R}$$

Brownian motion B_t and the compensated Poisson term $\tilde{N}(t, dz)$ are driving the stochastic differential equation in (23), and it is hence natural to formalize information with respect to these objects. We therefore let $\mathcal{F}_t$ denote the σ-algebra generated by B_s and $\tilde{N}(s, dz)$, $0 \le s \le t$. Intuitively $\mathcal{F}_t$ contains all the information up to time t. When information is delayed, we instead consider the σ-algebras

$$\mathcal{E}_t := \mathcal{F}_{t-\delta} \qquad t \in [\delta, T] \tag{24}$$

Both the retailer and the manufacturer should base their actions on the delayed information. Technically that means that q_t and w_t should be $\mathcal{E}_t$-adapted, i.e., q and w should be $\mathcal{E}$-predictable processes. In principle, the retail price R_t can be settled at a later stage. This case is possible to handle, but leads to complicated notation. We hence only consider the case where R_t is $\mathcal{E}$-predictable.

3.2. Finding Stackelberg equilibria in the newsvendor model

We now apply our general result for stochastic Stackelberg games to the newsvendor problem. In the newsvendor problem, we have the control $u = (u_L, u_F)$ where $u_L = w$ is the wholesale price, and $u_F = (q, R)$ with q the order rate and R the retail price. Moreover $X_t = D_t$,

$$f_L(t, X(t), u(t)) = (w_t - M)q_t, \qquad g_L = 0, \tag{25}$$

$$f_F(t, X(t), u(t)) = (R_t - S)\min(D_t, q_t) - (w_t - S)q_t, \qquad \text{and} \qquad g_F = 0. \tag{26}$$

Therefore by (10)

$$\begin{aligned} H_F(t, D_t, q_t, R_t, w_t, a_F(t), b_F(t), c_F(t,\cdot)) = (R_t - S)\min(D_t, q_t) - (w_t - S)q_t \\ + a_F(t)\mu(t, D_t, R_t) + b_F(t)\sigma(t, D_t, R_t) \\ + \int_{\mathbb{R}} \gamma(t, D_t, R_t, \xi)c_F(\xi)\nu(d\xi) \end{aligned} \tag{27}$$

Similarly by (14) , with $u_F = \phi(u_L) = (\phi_1(w), \phi_2(w)) = (q(w), R(w))$,

$$H_L^{\phi}(t, D_t, w_t, a_L(t), b_L(t), c_L(t,\cdot)) \tag{28}$$

$$=(w_t - M)\phi_1(t, w(t)) + a_L(t)\mu(t, D_t, \phi_2(t, w(t))) + b_L(t)\sigma(t, D_t, \phi_2(t, w(t)))) \tag{29}$$

$$+ \int_{\mathbb{R}} c_L(t,\xi)\gamma(t, D_t, \phi_2(t, w(t)), \xi)\nu(d\xi) \tag{30}$$

Here we have assumed that the dynamics of D_t only depends on the control $R_t = \phi_2(t, w(t))$ and has the general form

$$\begin{aligned} dD_t &= \mu(t, D_t, R_t)dt + \sigma(t, D_t, R_t)dB_t \\ &\quad + \int_{\mathbb{R}} \gamma(t, D_{t^-}, R_t, \xi)\tilde{N}(dt, d\xi); \qquad t \in [0, T] \end{aligned} \tag{31}$$

$$D_0 = d_0 \in \mathbb{R} \tag{32}$$

where $\mu(t, D, R), \sigma(t, D, R)$ and $\gamma(t, D, T, \xi)$ are continuous with respect to t and continuously differentiable (C^1) with respect to D and R. We choose $\mathcal{A}_{L,\mathcal{E}}, \mathcal{A}_{F,\mathcal{E}}$ to be the set of all $\mathcal{E}$-predictable processes with values in $\mathbb{U}_L = \mathbb{R}$ and $\mathbb{U}_F = \mathbb{R}^2$ respectively, where $\mathcal{E}_t = \mathcal{F}_{t-\delta}$ as above. Then we see that assumptions

(A1)-(A3) hold, with $\xi_L(t)$ and $\xi_F(t)$ given by $\xi_L(t) = 0$ for all $t \in [0,T]$ and

$$d\xi_F(t) = \xi_F(t)\left(\frac{\partial\mu}{\partial D}(t,D_t,R_t)dt + \frac{\partial\sigma}{\partial D}(t,D_t,R_t)dB_t + \int_{\mathbb{R}} \frac{\partial\gamma}{\partial D}(t,D_{t^-},R_t,\xi)\tilde{N}(dt,d\xi)\right) \tag{33}$$

$$+ \beta_F(t)\cdot\left(\frac{\partial\mu}{\partial R}(t,D_t,R_t)dt + \frac{\partial\sigma}{\partial R}(t,D_t,R_t)dB_t + \int_{\mathbb{R}} \frac{\partial\gamma}{\partial R}(t,D_{t^-},R_t,\xi)\tilde{N}(dt,d\xi)\right); t \in [0,T]$$

$$\xi_F(0) = 0 \tag{34}$$

where $\cdot$ denotes a vector product. To find a Stackelberg equilibrium we use Theorem 2.1. Hence by (19) we get the following first order conditions for the optimal values $\hat{q}_t = \phi_1(t,\hat{w}(t)), \hat{R}_t = \phi_2(t,\hat{w}(t))$:

$$\mathrm{E}\left[(\hat{R}_t - S)\mathcal{X}_{[0,D_t]}(\hat{q}_t) - w_t + S\middle|\mathcal{E}_t\right] = 0 \tag{35}$$

and

$$\mathrm{E}\Bigg[\min(D_t,\hat{q}_t) + a_F(t)\frac{\partial\mu}{\partial R}(t,D_t,\hat{R}) \tag{36}$$

$$+b_F(t)\frac{\partial\sigma}{\partial R}(t,D_t,\hat{R}) + \int_{\mathbb{R}} c_F(t,\xi)\frac{\partial\gamma}{\partial R}(t,D_t,\hat{R},\xi)\nu(d\xi)\Bigg|\mathcal{E}_t\Bigg] = 0$$

Here $\mathcal{X}_{[0,D_t]}(\hat{q}_t)$ denotes the indicator function for the interval $[0,D_t]$, i.e., a function that has the value 1 if $0 \le \hat{q}_t \le D_t$, and is zero otherwise. Let $\hat{q}_t = \Phi_1(w)(t), \hat{R}_t = \Phi_2(w)(t)$ be the solution of this coupled system. Next, by (20) we get the first-order condition

$$(\hat{w}_t - M)\phi_1'(t,\hat{w}(t)) + \phi_1(t,\hat{w}(t)) + \phi_2'(t,\hat{w}(t))\mathrm{E}\Bigg[a_L(t)\frac{\partial\mu}{\partial R}(t,D_t\phi_2(t,\hat{w}(t))) \tag{37}$$

$$+b_L(t)\frac{\partial\sigma}{\partial R}(t,D_t,\phi_2(t,\hat{w}(t))) + \int_{\mathbb{R}} c_L(t,\xi)\frac{\partial\gamma}{\partial R}(t,D_t,\phi_2(t,\hat{w}(t)),\xi)\nu(d\xi)\Bigg|\mathcal{E}_t\Bigg] = 0$$

for the optimal value $\hat{w}_t$. We summarize what we have proved in the following theorem.

Theorem 3.2.1

Suppose u^* is a Stackelberg equilibrium for the newsvendor problem with state $X_t = D_t$ given by (31) and performance functionals

$$J_L(w,(q,R)) = \mathrm{E}\left[\int_\delta^T (w_t - M)q_t dt\right] \qquad \textit{(manufacturer's profit)} \tag{38}$$

$$J_F(w,(q,R)) = \mathrm{E}\left[\int_\delta^T \Big((R_t - S)\min(D_t,q_t) - (w_t - S)q_t\Big)dt\right] \qquad \textit{(retailer's profit)} \tag{39}$$

Assume that (13), (A1)-(A3), and (17) hold. Let $\hat{q}_t = \phi_1(t, w(t)), \hat{R}_t = \phi_2(t, w(t))$ *be the solution of (35)–(36). Assume that* $\phi_i \in C^1$ *and that the conditions of Theorem 2.1 are satisfied. Let* $\hat{w}_t$ *be the solution of (37). Then*

$$u^* = (\hat{w}_t, (\phi_1(t, \hat{w}(t)), \phi_2(t, \hat{w}(t))) \in \mathcal{A}_{\mathcal{E}}$$

In other words

$$\max_{(q,R)\in\mathcal{A}_{F,\mathcal{E}}} \{J_F(w, (q, R))\} = J_F(w, (\Phi_1(w), \Phi_2(w))) \tag{40}$$

and

$$\max_{w\in\mathcal{A}_{L,\mathcal{E}}} \{J_L(w, (\Phi_1(w), \Phi_2(w)))\} = J_L(\hat{w}, (\Phi_1(\hat{w}), \Phi_2(\hat{w}))) \tag{41}$$

Remark

Note that if R is fixed and cannot be chosen by the retailer, then (36) is irrelevant and we are left with (35) leading to the simpler equations in Theorem 3.2.2. In the special case when D_t does not depend on R_t, we get:

Theorem 3.2.2

Assume that D_t *has a continuous distribution, that* D_t *does not depend on* R_t *and that* $R_t = R$ *is exogenously given and fixed. For any given* w_t *with* $S < M \le w_t \le R$ *consider the equation*

$$\mathrm{E}\left[(R - S)\mathcal{X}_{[0,D_t]}(q_t) - w_t + S|\mathcal{E}_t\right] = 0 \tag{42}$$

Let $q_t = \phi_1(t, w(t))$ *denote the unique solution of (42), and assume that the function*

$$w_t \mapsto \mathrm{E}\left[(w_t - M)\phi_1(t, w(t))\right] \tag{43}$$

has a unique maximum at $w_t = \hat{w}_t$*. Then with* $\hat{q}_t = \phi_1(t, \hat{w}(t))$ *the pair* $(\hat{w}, \hat{q})$ *is a unique Stackelberg equilibrium for the newsvendor problem defined by (22) and (21).*

Proof

To see why (42) always has a unique solution, note that w_t is $\mathcal{E}_t$-measurable and hence (42) is equivalent to

$$\mathrm{E}\left[\mathcal{X}_{[0,D_t]}(q_t)|\mathcal{E}_t\right] = \frac{w_t - S}{R - S} \tag{44}$$

Existence and uniqueness of q_t then follows from monotonicity of conditional expectation. Uniqueness of the Stackelberg equilibrium follows from Theorem 3.2.1. To see that the candidate $\hat{q}_t = \Phi_1(\hat{w})(t)$ is indeed a Stackelberg Equilibrium, we argue as follows: Since the maximum $\hat{w}_t$ is unique, any other w_t will lead to strictly lower expected profit at time t. As demand does not depend on w_t, low expected profit at one point in time cannot be compensated by higher expected profits later on. Hence if the statement $w_t = \hat{w}_t$ a.s. $\lambda \times P$ (λ denotes Lebesgue measure on $[0, T]$) is false, any such strategy will lead to strictly lower expected profits. The same argument applies for the retailer, and hence $(\hat{w}, \hat{q})$ is a Stackelberg equilibrium.

□

To avoid degenerate cases we need to know that D_t has a continuous distribution. In the next sections we will consider special cases, and we will often be able to write down explicit solutions to (42) and prove that (43) has a unique maximum. Notice that (42) is an equation defined in terms of conditional expectation. Conditional statements of this type are in general difficult to compute, and the challenge is to state the result in terms of unconditional expectations.

4. Explicit solution formulas

In this section we will assume that the conditions of Theorem 2.1 hold.

4.1. The Ornstein-Uhlenbeck process with constant coefficients

In this section, we offer explicit formulas for the equilibria that occur when the demand rate is given by a constant coefficient Ornstein–Uhlenbeck process, i.e., the case where D_t is given by

$$dD_t = a(\mu - D_t)dt + \sigma dB_t \tag{45}$$

where a, μ, and σ are constants. The Ornstein–Uhlenbeck process is important in many applications. In particular, it is commonly used as a model for the electricity market. The process is *mean reverting* around the constant level μ, and the constant a decides the speed of mean reversion. The explicit solution to (45) is

$$D_t = D_0 e^{-at} + \mu(1 - e^{-at}) + \int_0^t \sigma e^{a(s-t)} dB_s \tag{46}$$

It is easy to see that

$$D_t = D_{t-\delta} e^{-a\delta} + \mu(1 - e^{-a\delta}) + \int_{t-\delta}^t \sigma e^{a(s-t)} dB_s \tag{47}$$

Because the last term is independent of $\mathcal{E}_t$ with a normal distribution $N(0, \frac{\sigma^2(1-e^{-2a\delta})}{2a})$, it is easy to find a closed form solution to (42). We let $G[z]$ denote the cumulative distribution of a standard normal distribution, and $G^{-1}[z]$ its inverse. The final result can be stated as follows:

Proposition 4.1.1

For each $y \in \mathbb{R}$, let $\Phi_y : [M, R] \to \mathbb{R}$ denote the function

$$\Phi_y[w] = ye^{-a\delta} + \mu(1 - e^{-a\delta}) + \sigma\sqrt{\frac{1 - e^{-2a\delta}}{2a}} \cdot G^{-1}\left[1 - \frac{w - S}{R - S}\right] \tag{48}$$

and let $\Psi_y : [M, R] \to \mathbb{R}$ denote the function $\Psi_y[w] = (w - M)\Phi_y[w]$. If $\Phi_y[M] > 0$, the function Ψ_y is quasiconcave and has a unique maximum with a strictly positive function value.

At time $t - \delta$ the parties should observe $y = D_{t-\delta}$, and a unique Stackelberg equilibrium is obtained at

$$w_t^* = \begin{cases} \text{Argmax}[\Psi_y] & \text{if } \Phi_y[M] > 0 \\ M & \text{otherwise} \end{cases} \qquad q_t^* = \begin{cases} \Phi_y[\text{Argmax}[\Psi_y]] & \text{if } \Phi_y[M] > 0 \\ 0 & \text{otherwise} \end{cases} \tag{49}$$

To prove Proposition 4.1.1, we need the following lemma.

Lemma 4.1.2

In this lemma $G[x]$ is the cumulative distribution function of the standard normal distribution. Let $0 \le m \le 1$, and for each m consider the function $h_m : \mathbb{R} \to \mathbb{R}$ defined by

$$h_m[z] = z(1 - m - G[z]) - G'[z] \tag{50}$$

Then

$$h_m[z] < 0 \qquad \textit{for all } z \in \mathbb{R} \tag{51}$$

Proof of Lemma 4.1.2

Note that if $z \ge 0$, then $h_m[z] \le h_0[z]$ and if $z \le 0$, then $h_m[z] \le h_L[z]$. It hence suffices to prove the lemma for $m = 0$ and $m = 1$. Using $G''[z] = -z \cdot G'[z]$, it is easy to see that $h''_m[z] = -G'[z] \le 0$. If $m = 0$, it is straightforward to check that h_0 is strictly increasing, and that $\lim_{z\to+\infty} h_0[z] = 0$. If $m = 1$, it is straightforward to check that $h_L[z]$ is strictly decreasing, and that $\lim_{z\to-\infty} h_L[z] = 0$. This proves that h_0 and h_L are strictly negative, completing the proof of the lemma.

□

Proof of Proposition 4.1.1

From (47), we easily see that the statement $q_t \leq D_t$ is equivalent to the inequality

$$q_t - \left(D_{t-\delta}e^{-a\delta} + \mu(1-e^{-a\delta}\right) \leq \int_{t-\delta}^{t} \sigma e^{a(s-t)} dB_s \tag{52}$$

The left-hand side is $\mathcal{E}_t$-measurable, while the right-hand side is normally distributed and independent of $\mathcal{E}_t$. Using the Itô isometry, we see that the right-hand side has expected value zero and variance $\frac{\sigma^2(1-e^{-2a\delta})}{2a}$. It is then straightforward to see that

$$\mathrm{E}\left[\mathcal{X}_{[0,D_t]}(\hat{q}_t)|\mathcal{E}_t\right] = 1 - G\left[\frac{q_t - \left(D_{t-\delta}e^{-a\delta} + \mu(1-e^{-a\delta})\right)}{\sqrt{\frac{\sigma^2(1-e^{-2a\delta})}{2a}}}\right] \tag{53}$$

and (48) follows trivially from (44). It remains to prove that the function Ψ_y has a unique maximum if $\Phi_y[M] > 0$. First put

$$\hat{y} = \frac{y \cdot e^{-a\delta} + \mu(1-e^{-a\delta})}{\sigma\sqrt{\frac{1-e^{-2a\delta}}{2a}}} \tag{54}$$

and note that Ψ_y is proportional to

$$(w - M)\left(\hat{y} + G^{-1}\left[1 - \frac{w-S}{R-S}\right]\right) \tag{55}$$

We make a monotone change of variables using $z = G^{-1}\left[1 - \frac{w-S}{R-S}\right]$. With this change of variables we see that Ψ_y is proportional to

$$(R-S)\left(1 - G[z] - \frac{M-S}{R-S}\right)(\hat{y} + z) \tag{56}$$

Put $m = \frac{M-S}{R-S}$, and note that Ψ_y is proportional to

$$(1 - m - G[z])(\hat{y} + z) \tag{57}$$

$\Phi_y[M] > 0$ is equivalent to $\hat{y} + G^{-1}[1-m] > 0$, and the condition $w \geq M$ is equivalent to $z \leq G^{-1}[1-m]$. Note that if $S \leq M \leq R$, then $0 \leq m \leq 1$. For each fixed $0 \leq m \leq 1, \hat{y} \in \mathbb{R}$ consider the function

$$f_m[z] = (1 - m - G[z])(\hat{y} + z) \qquad \text{on the interval } -\hat{y} \leq z \leq G^{-1}[1-m] \tag{58}$$

If $\hat{y} + G^{-1}[1-m] > 0$, the interval is nondegenerate and nonempty, and

$$f_m'[z] = -G'[z](\hat{y} + z) + (1 - m - G[z]) \tag{59}$$

Note that $f_m'[-\hat{y}] > 0$, and that $f_m[-\hat{y}] = f_m[G^{-1}[1-m]] = 0$. These functions therefore have at least one strictly positive maximum. To prove that the maximum is unique, assume that $f_m'[z_0] = 0$, and compute $f_m''[z_0]$. Using $G''[z] = -z \cdot g'[z]$, it follows that

$$f_m''[z_0] = z_0(1 - m - G[z_0]) - 2G'[z_0] < z_0(1 - m - G[z_0]) - G'[z_0] < 0 \tag{60}$$

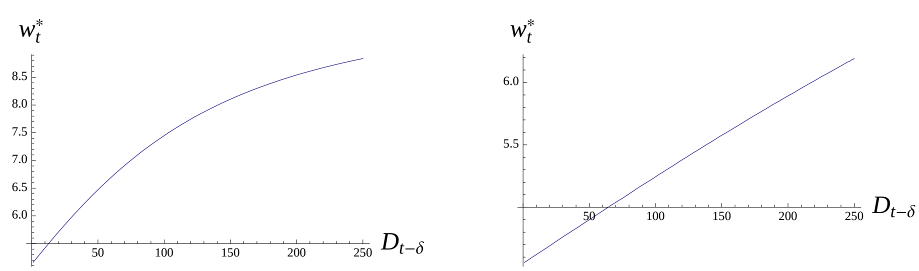
by Lemma 4.1.2. The function is thus quasiconcave and has a unique, strictly positive maximum. Existence of a unique Stackelberg equilibrium then follows from Theorem 3.2.2.

□

The condition $\Phi_y[M] > 0$ has an obvious interpretation. The manufacturer cannot offer a wholesale price w lower than the production cost M. If $\Phi_y[M] \leq 0$, it means that the retailer is unable to make a positive expected profit even at the lowest wholesale price the manufacturer can offer. When that occurs, the retailer's best strategy is to order $q = 0$ units. When the retailer orders $q = 0$ units, the choice of w is arbitrary. However, the choice $w = M$ is the only strategy that is increasing and continuous in y.

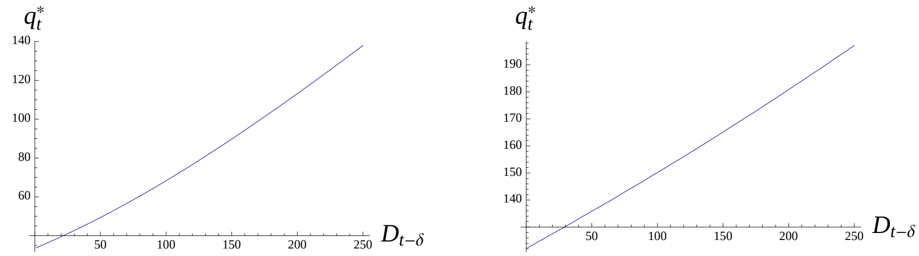
Given values for the parameters a, μ, σ, S, M, R, and δ, the explicit expression in (48) makes it straightforward to construct the deterministic function $y \mapsto \text{Argmax}[\Psi_y]$ numerically. Two different graphs of this function are shown in Figure 2. Figure 3 shows the corresponding function $\Phi_y[\text{Argmax}[\Psi_y]]$. In the construction we used a delay $\delta = 7$ and $\delta = 30$, with the parameter values

$$a = 0.05 \qquad \mu = 100 \qquad \sigma = 12 \qquad R = 10 \qquad S = 1 \qquad M = 2 \tag{61}$$



$\delta = 7$ $\qquad$ $\delta = 30$

Figure 2: w_t^* as a function of the observed demand rate $D = D_{t-\delta}$



$\delta = 7$ $\delta = 30$

Figure 3: q_t^* as a function of the observed demand rate $D = D_{t-\delta}$

Note that the manufacturing cost $M = 2$ is relatively low, and $\Phi_y[M] > 0$ is satisfied for all $y > 0$ in these cases. It is interesting to note that the equilibria change considerably when the delay increases from $\delta = 7$ to $\delta = 30$ (notice the scale on the y-axis).

4.2. Further applications to the case with fixed R

Explicit results like the one in Proposition 4.1.1 can be carried out in a number of different cases. A complete discussion of these cases would be too long to be include here, and is provided in a separate paper Øksendal et al (2012). To demonstrate the usefulness of this theory, we briefly survey the results in Øksendal et al (2012):

- The Ornstein-Uhlenbeck process with time-variable (deterministic) coefficients: Existence, uniqueness, and explicit solutions for the equilibria.

- Geometric Brownian motion with constant coefficients: Existence, uniqueness, and explicit solutions for the equilibria. Interestingly, the equilibrium wholesale price w_t is constant in this case, and the retailer orders a fixed fraction of the observed demand.

- Geometric Brownian motion with time-variable (deterministic) coefficients: Existence, uniqueness, and explicit solutions for the equilibria. In this case the equilibrium wholesale price is not constant. It is, however, given by a deterministic function, and as a consequence the manufacturer needs not observe demand to settle the price.

- Geometric Lévy processes: Explicit solutions for the equilibria are offered for special cases with random coefficients, leading to non-Markovian solutions. Existence and uniqueness is established for some cases. Typically the manufacturer has an equilibrium wholesale price defined in terms of a deterministic function, and needs not observed demand. The retailer should observe both demand and the growth rate of demand as his optimal order q is a deterministic function of these two quantities.

5. *R*-dependent demand

In this section we provide a solution to an example with R-dependent demand. This problem is more difficult than the case we handled in the previous section. We also discuss a more complicated example, raising some interesting issues for future research.

5.1. An example with R-dependent demand

When demand depends on R_t, Theorem 3.2.2 no longer applies. High profits at some stage may become too costly later, due to reduced demand, and the problem can no longer be separated into independent one-periodic problems. In particular we shall see that (13) no longer holds. However, we can still apply the maximum principle for the optimization of J_F (the follower problem), since this part does not need (13). To simplify the discussion, we note that in the particular case where the coefficients μ, σ, γ do not depend on D, then the adjoint equations (15)–(16) have the trivial solution $a_L(t) = b_L(t) = 0$. If

$$dD_t = (K - R_t)dt + \sigma dB_t \tag{62}$$

the second pair of adjoint variables is also solvable, i.e., (11)–(12) has the explicit solution

$$a_F(t) = \mathrm{E}\left[\int_t^T (R_s - S)\mathcal{X}_{[0,q_s]}(D_s)ds|\mathcal{E}_t\right] \qquad b_F(t) = 0 \tag{63}$$

If we make the simplifying assumption that R_t is decided at time $t-\delta$, i.e., at the same time as w_t and q_t, then, maximizing the Hamiltonian H_F as in Theorem 3.2.1, we arrive by (33) and (34) at the following first-order conditions for the optimal functions $w_t = \hat{w}_t, q = \hat{q}_t = \Phi_1(\hat{w})(t)$ and $R = \hat{R}_t = \Phi_2(\hat{w})(t)$:

$$\mathrm{E}\left[\mathcal{X}_{[0,D_t^+]}(q_t)|\mathcal{E}_t\right] = \frac{w_t - S}{R_t - S} \qquad t \in [\delta, T] \tag{64}$$

$$\mathrm{E}\left[\min[D_t, q_t] - \int_t^T (R_s - S)\mathcal{X}_{[0,q_s]}(D_s)ds|\mathcal{E}_t\right] = 0 \qquad t \in [\delta, T] \tag{65}$$

The functions ϕ_1 and ϕ_2 are found solving (64)-(65), as explained in the following.

It is interesting to note that while (64) can be solved pointwise in t, (65) depends on path properties in the remaining time period, reflecting that decisions taken at one point in time influence later performance.

The optimal order quantity $q_t = \Phi_1(w)(t)$ can be found from the equations as follows: Using the same separation technique that we used in Section 4, we can express q_t explicitly in terms of w_t and R_t:

$$q_t = D_{t-\delta} + \int_{t-\delta}^t (K - R_s)ds + \sqrt{\sigma\delta} \cdot G^{-1}\left[1 - \frac{w_t - S}{R_t - S}\right] \tag{66}$$

If we put $t = \delta$, we obtain

$$q_\delta = D_0 + \int_0^\delta (K - R_s)ds + \sqrt{\sigma\delta} \cdot G^{-1}\left[1 - \frac{w_\delta - S}{R_\delta - S}\right] \tag{67}$$

The interesting point here is that we need to know the prices $R_t, 0 \leq t \leq \delta$ in the period prior to the sales period $[\delta, T]$. One option is to consider these values as exogenously given initial values, which is typical when handling differential equations with delay. Alternatively, these prior values can be considered part of the decision process. In that case, the choice $R_t = 0$ if $0 \leq t \leq \delta$ is optimal as it leads to higher values of initial demand, clearly an advantage for both the retailer and the manufacturer. This strategy corresponds to advertising in the presales period, in which case a small number of items are given away free to stimulate demand.

We now proceed to solve (64)–(65): By (64) we obtain

$$\mathrm{E}\left[\mathcal{X}_{[0,q_t]}(D_t)|\mathcal{E}_t\right] = 1 - \frac{w_t - S}{R_t - S} \tag{68}$$

The function

$$x \mapsto h_t(x) := \mathrm{E}[\mathcal{X}_{[0,x]}(D_t)|\mathcal{E}_t]$$

is strictly increasing and hence has an inverse $h_t^{-1}(x)$. Thus (68) can be written

$$\begin{aligned} q_t &= h_t^{-1}\left(\frac{R_t - w_t}{(R_t - w_t) + w_t - S}\right) \\ &= h_t^{-1}\left(\frac{y}{y + w_t - S}\right)_{y = R_t - w_t} \end{aligned} \tag{69}$$

If we substitute (68) into (65), we get

$$\mathrm{E}\left[\int_t^T (R_s - S)\left(1 - \frac{w_s - S}{R_s - S}\right) ds \Big| \mathcal{E}_t\right] = \mathrm{E}[\min[D_t, q_t]|\mathcal{E}_t],$$

or

$$\mathrm{E}\left[\int_t^T (R_s - w_s)ds \Big| \mathcal{E}_t\right] = Y_t, \tag{70}$$

where

$$\begin{aligned} Y :&= \mathrm{E}[\min[D_t, q_t]|\mathcal{E}_t] = \mathrm{E}[q_t \mathcal{X}_{[0,D_t]}(q_t)|\mathcal{E}_t] + f_t(q_t) \\ &= q_t \frac{w_t - S}{R_t - S} + f_t(q_t), \end{aligned} \tag{71}$$

with

$$f_t(x) = \mathrm{E}[D_t \mathcal{X}_{[0,x]}(D_t)|\mathcal{E}_t]. \tag{72}$$

Hence, by (69)

$$Y_t = F_t(w_t, R_t - w_t), \tag{73}$$

where

$$F_t(w,y) = h_t^{-1}\left(\frac{y}{y+w-S}\right)\frac{w-S}{y+w-S} + f_t\left(h_t^{-1}\left(\frac{y}{y+w-S}\right)\right). \tag{74}$$

For each fixed t and w, let $F_t^{-1}(w,\cdot)$ be a measurable left-inverse of the mapping

$$y \mapsto F_t(w,y),$$

in the sense that

$$F_t^{-1}(w, F_t(w,y)) = y \qquad \text{for all } y \in \mathbb{R}. \tag{75}$$

Then

$$R_s - w_s = F_s^{-1}(w_s, Y_s);\ \ s \in [0,T]. \tag{76}$$

Therefore equation (70) can be written

$$\mathrm{E}\left[\int_t^T F_s^{-1}(w_s,Y_s)ds\Big|\mathcal{E}_t\right] = Y_t; \qquad t \in [\delta, T] \tag{77}$$

This is a backward stochastic differential equation (BSDE) in the unknown process Y_t. It can be reformulated as follows: Find an $\mathcal{E}_t$-adapted process Y_t and an $\mathcal{E}_t$-martingale Z_t such that

$$\begin{cases} dY_t &= -F_t^{-1}(w_t,Y_t)dt + dZ_t; \qquad t \in [\delta, T] \\ Y_T &= 0 \end{cases} \tag{78}$$

From known BSDE theory we obtain the existence and uniqueness of a solution for (Y_t, Z_t) of such an equation under certain conditions on the driver process $F_t^{-1}(w_t,Y_t)$. For example, it suffices that

$$\mathrm{E}\left[\int_\delta^T F_t^{-1}(w_t,0)^2 dt\right] < \infty \qquad \text{and} \qquad y \mapsto F_t^{-1}(w_t,y) \text{ is Lipschitz} \tag{79}$$

See, e.g., Pardoux and Peng (1990) or El Karoui et al. (1997) and the references therein. Moreover, Y_t and

Z_t can be obtained as a fixed point of a contraction operator and hence as a limit of an iterative procedure. This makes it possible to compute Y_t numerically in some cases. In general, however, the solution of the BSDE (78) need not be unique, because $F_t^{-1}(w_t,\cdot)$ is not necessarily unique, and, even if F_t is invertible it is not clear that the inverse satisfies (79). If we assume that the solution $Y_t = Y_t(\omega)$ of (78) has been found, then the optimal $R_t = \hat{R}_t(w) = \Phi_2(w)$ is given by (76) and the optimal $q_t = \hat{q}_t(w) = \Phi_1(w)(t)$ is given by (69).

Finally we turn to the manufacturer´s maximization problem. The performance functional for the manufacturer has the form:

$$J_L(w, \Phi(w)) = \mathrm{E}\left[\int_0^T (w_t - M)(\hat{q}(w))_t dt\right] \tag{80}$$

Therefore, by (69) and by (76) the problem to maximize $J_L(w, \Phi(w))$ over all paths w, can be regarded as a problem of optimal stochastic control of a coupled system of forward-backward stochastic differential equations (FBSDEs), as follows:

(Forward system)

$$dD_t = (K - R_t)dt + \sigma dB_t = (K - w_t - F_t^{-1}(w_t, Y_t))dt + \sigma dB_t;\ D_0 \in \mathbb{R} \tag{81}$$

(Backward system)

$$\begin{cases} dY_t &= -F_t^{-1}(w_t, Y_t)dt + dZ_t; \qquad t \in [\delta, T] \\ Y_T &= 0 \end{cases} \tag{82}$$

(Performance functional)

$$J(w) := \mathrm{E}\left[\int_0^T (w_t - M)h_t^{-1}\left(\frac{F_t^{-1}(w_t, Y_t)}{F_t^{-1}(w_t, Y_t) + w_t - S}\right) dt\right] \tag{83}$$

This is a special case of the following stochastic control problem of a coupled system of FBSDEs:

(Forward system)

$$dX_t = b(t, X_t, Y_t, u_t)dt + \sigma(t, X_t, Y_t, u_t)dB_t \tag{84}$$

$$X_0 \in \mathbb{R}$$

(Backward system)

$$dY_t = -g(t, u_t, Y_t)dt + dZ_t \tag{85}$$
$$Y_T = 0$$

(Performance functional)

$$J(u) = \mathrm{E}\left[\int_0^T f(t, X_t, Y_t, u_t)dt\right] \tag{86}$$

Here u_t is our control. To handle this problem, we need an extension of the result in Øksendal and Sulem (2012) to systems with the coupling given in (84) and (86). The extension is straightforward and we get the following solution procedure:

Define the Hamiltonian

$$H(t, x, y, w, \lambda, p, q) = (w - M)\hat{q}_t(w - y) + \lambda F_t^{-1}(w, y) + p(K - w - F_t^{-1}(w, y)) + q\,\sigma \tag{87}$$

The adjoint processes λ_t, p_t, q_t are given by the following FB system:

$$d\lambda_t = \frac{\partial H}{\partial y}(t)dt = \left((w_t - M)(-\hat{q}_t'(w_t - Y_t) + \lambda_t \frac{d}{dy}\left(F_t^{-1}(w_t, y)\right)_{y=Y_t}\right) dt\ ;\ t \geq 0 \tag{88}$$
$$\lambda_0 = 0$$

$$dp_t = -\frac{\partial H}{\partial x}(t)dt + q_t dB_t = q_t dB_t \tag{89}$$
$$p_T = 0$$

From (89) we get $p_t = q_t = 0$, and the first order condition for maximization of the functional $w \mapsto H(t, D_t, w, \lambda_t, p_t, q_t)$ becomes

$$(\hat{w} - M)\hat{q}_t'(w_t - Y_t) + \hat{q}_t(w_t - Y_t) + \hat{\lambda}(t)\frac{d}{dw}\left(F_t^{-1}(w, \hat{Y}(t)\right)_{w=\hat{w}_t} = 0 \tag{90}$$

We formulate what we have proved in a proposition.

Proposition 5.1.1

Suppose that the demand process is as in (62) and that $\mathcal{E}_t = \mathcal{F}_{t-\delta};\ t \geq \delta$. Suppose that an optimal *solution* $\hat{w}_t, \hat{q}_t = \Phi_1(\hat{w})(t)$*, and* $\hat{R}_t = \Phi_2(\hat{w})(t)$ *of the Stackelberg game (21)–(22) exists. Then the retailer's optimal order response* $q_t = \Phi_1(w)(t)$ *and optimal price* $R_t = \Phi_2(w)(t)$*, respectively, are given by*

$$\Phi_1(w)(t) = h_t^{-1}\left(\frac{R_t - w_t}{R_t - S}\right) \tag{91}$$

$$\Phi_2(w)(t) = w_t + F_t^{-1}(w_t, Y_t), \tag{92}$$

where $Y_t = Y_t^{(w_t)}$ *is a solution of the BSDE (78) for some measurable left inverse* $F_t^{-1}(w_t,\cdot)$ *of* $F_t(w_t,\cdot)$*. Accordingly, the manufacturer's wholesale price* $\hat{w}_t$ *is the solution* w_t *of equation (90).*

Some remarks

Even though the result in Proposition 5.1.1 only covers a special case, we believe that the solution features insights to more general cases. We see that once R_t is decided, the order quantity q_t can be found via a pointwise optimization. This is true because the order size does not influence demand, and a suboptimal choice at time t cannot be compensated by improved performance later on. We expect this strategy to hold more generally.

Once q_t is eliminated from the equations, the optimal retail price is found via a transformation to a backward stochastic differential equation. We believe that similar transformations might work for other cases. It makes good sense that the optimal retail price satisfies a backward problem. As we approach time T, it becomes less important what happens later on. In the limiting stages we take all we can get, leading to an end-point constraint.

If F_t is not invertible, our framework will allow for solutions that might jump to new levels. Solutions of this type are found regularly when solving ordinary stochastic control problems. Our setup appears to allow for a similar type of effect in a quite unexpected way. A possible conjecture is that there exist switching levels, i.e., when demand reaches a low level the retailer should stop selling and lower prices to increase demand (sell marginal quantities with marketing effects in mind), and start selling when demand reaches a high enough level. Non-uniqueness of F_t^{-1} could lead to switching effects of this kind. This is an interesting problem which is left for future research.

5.2. A second example allowing complete elimination of the adjoint equations

Another model admitting a similar type of analysis is:

$$dD_t = D_t(K - R_t)dt + \sigma D_t dB_t \tag{93}$$

This is a second example where the adjoint equations can be solved explicitly, eventually leading to a system of the form

$$\mathrm{E}\left[\mathcal{X}_{[0,D_t]}(q_t)|\mathcal{E}_t\right] = \frac{w_t - S}{R_t - S} \qquad t \in [\delta, T] \tag{94}$$

$$\mathrm{E}\left[\min[D_t, q_t] - \frac{D_t}{\Gamma_F(t)} \cdot \int_t^T (R_s - S)\mathcal{X}_{[0,q_s]}(D_s)\Gamma_F(s)ds|\mathcal{E}_t\right] = 0 \qquad t \in [\delta, T] \tag{95}$$

$$(w_t - M)\phi_L{}'(w_t) + \phi_L(w_t) - \phi_F{}'(w_t) \cdot \mathrm{E}\left[\frac{D_t}{\Gamma_L(t)}\int_t^T \Gamma_L(s)ds|\mathcal{E}_t\right] = 0 \qquad t \in [\delta, T] \tag{96}$$

where

$$d\Gamma_L(t) = \Gamma_L(t)(-\phi_F(w_t)dt + \sigma b_L(t)dB_t) \qquad \Gamma_L(0) = 1 \tag{97}$$

$$d\Gamma_F(t) = \Gamma_F(t)(R_t dt + \sigma b_F(t)dB_t) \qquad \Gamma_F(0) = 1 \tag{98}$$

We see that even though the adjoint equations can be eliminated, the resulting system is an order of magnitude more complicated than (64)–(**??**). We have not been able to find a solution to this case. More refined solution procedures that could handle such problems analytically or numerically would be of great value, and is an interesting topic for future research.

5.3. An example with explicit solution

In this section we consider a simplified case with R-dependent demand, but where the contract must be written upfront, i.e., that w, q, R are decided once and for all prior to the sales period. This corresponds to the case when

$$\mathcal{E}_t = \mathcal{E}_t = \mathcal{F}_0 = \{\emptyset, \Omega\} \qquad \text{for all } t \in [\delta, T].$$

so that

$$\mathrm{E}[Y|\mathcal{E}_t] = \mathrm{E}[Y] \qquad \text{for all } t$$

It can be shown that the maximum principle can be modified to cover this situation. We do not give the proof here, but refer to the argument given in Section 10.4 in Øksendal and Sulem (2007): When the controls w, q, R are not allowed to depend on t, we consider the t-integrated Hamiltonians, given by (see (12)-(17))

$$\begin{aligned}\tilde{H}_F(w,q,R,a_F,b_F,c_F) :=& \int_\delta^T H_F(t,D_t,w,q,R,a_F(t),b_F(t),c_F(t))dt \\ =&(R-S)\int_\delta^T \min[D_t,q]dt - (w-S)qT + (K-R)\int_\delta^T a_F(t)dt \end{aligned} \tag{99}$$

and

$$\begin{aligned}\tilde{H}_L{}^\phi(w,a_L,b_L,c_L) :=& \int_\delta^T H_L{}^\phi(t,D_t,w,a_L(t),b_L(t),c_L(t))dt \\ =&(w-M)\phi_L(w)T = (w-M)\hat{q}(w)T, \end{aligned} \tag{100}$$

where $\phi(w) = (\phi_L(w), \phi_F(w)) = (\hat{q}(w), \hat{R}(w))$ is the maximizer with respect to q and R of

$$\begin{aligned}(p,q) \mapsto& \mathrm{E}[\tilde{H}(w,p,q,a_F,b_F,c_F)] \\ =&(R-S)\mathrm{E}\left[\int_\delta^T \min[D_t,q]dt\right] - (w-S)qT + (K-R)\mathrm{E}\left[\int_\delta^T a_F(t)dt\right] \end{aligned} \tag{101}$$

The optimal constant value $w = \hat{w}$ chosen by the manufacturer, is then the maximizer of

$$w \mapsto \mathrm{E}[\tilde{H}_L{}^\phi(w,a_L,b_L,c_L)] = (w-M)\hat{q}(w)T.$$

This gives the following first order conditions for the optimal (q, R) and w, (see (63)-(64))

$$\mathrm{E}\left[\int_\delta^T \mathcal{X}_{[0,D_t^+]}(q)dt\right] = \frac{w-S}{R-S}, \tag{102}$$

$$\mathrm{E}\left[\int_\delta^T \min[D_t,q]dt\right] = (R-S)\mathrm{E}\left[\int_\delta^T \left(\int_t^T \mathcal{X}_{[0,q]}(D_s)ds\right)dt\right] \tag{103}$$

and

$$(w-M)\frac{d}{dw}\hat{q}(w,\hat{R}(w)) + \hat{q}(w,\hat{R}(w)) = 0, \tag{104}$$

where $\hat{q}(w, R)$ is the solution of (90) for given w, R. Substituting $q = \hat{q}(w, R)$ into (91), we find the optimal $\hat{R} = \hat{R}(w)$ by solving (91), which can be reformulated to

$$\hat{R} = S + \frac{\mathrm{E}\left[\int_\delta^T \min[D_t, \hat{q}(w, \hat{R})]dt\right]}{\mathrm{E}\left[\int_\delta^T t\mathcal{X}_{[0,\hat{q}(w,\hat{R})]}(D_t)dt\right]} \tag{105}$$

by changing the order of integration. The main result can be summarized in the following theorem:

Theorem 5.3.1

If the values of q, R and w are required to be constant, then the optimal values $\hat{q}, \hat{R}, \hat{w}$ are given as follows: For given (w, R) let $\hat{q}(w, R)$ be the solution of (90). Next let $\hat{R} = \hat{R}(w)$ be the solution of (93). Then $\hat{w}$ is found as the solution of (92).

To get an impression how this works in a specific case, we consider the demand given by (62). In that example we have

$$P(D_t \leq q) = G\left[\frac{q - D_0 - (K - R)t}{\sigma\sqrt{t}}\right] \tag{106}$$

The equation (90) takes the form

$$\int_\delta^T P(D_t \leq q)dt = T - \frac{w - S}{R - S} \tag{107}$$

This equation clearly has no analytical solution, but can be handled numerically. (93) leads to the equation

$$R = S + \frac{\int_\delta^T \int_\delta^q \frac{z}{\sqrt{2\pi t\sigma^2}} e^{-\frac{(z - D_0 - (K-R)t)^2}{2\sigma\sqrt{t}}} dz + q \cdot (1 - P(D_t \leq q))dt}{\int_\delta^T t \cdot P(D_t \leq q)dt} \tag{108}$$

which we also think is reasonable to handle, as given w this is an equation in 1-variable only. Tables of $q = q(R, w)$ and $R = R(w)$ can then be constructed numerically, and values from these tables can be used to find maximum of the 1-variable function

$$w \mapsto (w - M)q(R(w), w)T$$

Since this process is Markov, we see that the parties only need to know D_0 to write the contract at time $t = 0$.

6. Concluding remarks

This paper has two main topics. First, we develop a new theory for stochastic differential Stackelberg games and second we apply that theory to continuous time newsvendor problems. In the continuous time newsvendor problem we offer a full description of the general case where our stochastic demand rate D_t is a function of the retail price R_t. The wholesale price w_t and the order rate q_t are decided based on information present at time $t-\delta$, while the retail price can in general be decided later, i.e., at time $t-\delta_R$ where $\delta > \delta_R$. This problem can be solved using Theorem 3.2.1. However, the solution is defined in terms of a coupled system of stochastic differential equations and in general such systems are hard to solve in terms of explicit expressions.

The case where demand is independent of R, leads to the simpler version in Theorem 3.2.2. If demand is given by an Ornstein-Uhlenbeck process, there is a unique, closed form solution to the problem. In Section 5 we have discussed some examples with R-dependent demand. These cases are simple, but nonetheless they appear to capture important economic effects. It would hence be quite interesting if one could solve such problems using more refined expressions. A further analysis of these and similar examples poses real challenges, however, and much more work will be needed before we can understand these issues in full. This is, therefore, an interesting topic for future research.

Acknowledgement

The authors wish to thank Steve LeRoy for several useful comments on the paper. The authors gratefully acknowledge several valuable suggestions from the referees.

References

[1] Baghery F. and B. Øksendal, 2007. A maximum principle for stochastic control with partial information. *Stoch. Anal. Appl.* 25, 705–717.

[2] Bensoussan A., M. Çakanyıldırım, and S. P. Sethi 2007. A multiperiod newsvendor problem with partially observed demand. *Mathematics of Operations Research* 23, 2, 322–344.

[3] Bensoussan A., M. Çakanyıldırım and S. P. Sethi 2009. Optimal ordering policies for stochastic inventory problems with observed information delays. *Production and Operations Management* 18, Issue 5, 546–559.

[4] Berling P., 2006. Real options valuation principle in the multi-period base-stock problem. *Omega* 36, 1086–1095.

[5] Cachón G. P., 2003. Supply chain coordination with contracts. In: *The Handbook of Operations Research and Management Science: Supply Chain Management: Design, Coordination and Operation.* Chapter 6. A. G. de Kok and S. C. Graves (eds), Amsterdam: Elsevier, pp 229–340.

[6] Calzolari A., P. Florchinger, and G. Nappo (2011). Nonlinear filtering for stochastic systems with fixed delay: Approximation by a modified Milstein scheme. *Computers and Mathematics with Applications* 61, 9, 2498–2509.

[7] El Karoui N., S. Peng, and M.-C. Quenez, 1997. Backward stochastic differential equations in finance. *Mathematical Finance* 7, 1–71.

[8] Framstad N., B. Øksendal, and A. Sulem, 2004. Sufficient stochastic maximum principle for optimal control of jump diffusions and applications to finance. *J. Opt. Theor. Appl.* 121, 77–98.

[9] Gallego G. and I. Moon, 1993. The distribution free newsboy problem: Review and extensions. *Journal of the Operational Research Society* 44, 8, 825–834.

[10] Karlin, S. and C.R. Carr, 1962. Prices and optimal inventory policy. In K. Arrow, S. Karlin, and H. Scarf (editors), *Studies in Applied Probability and Management Science*, Stanford University Press, Stanford, pp. 159–172.

[11] Kogan K. and S. Lou, 2003. Multi-stage newsboy problem: A dynamical model. *European Journal of Operational Research* 149, 2, 448–458.

[12] Kogan K., 2003. Scheduling parallel machines by the dynamic newboy problem. *Computers & Operations Research* 31, 3, 429–443.

[13] Lariviere, M. A. and E. L. Porteus, 2001. Selling to the Newsvendor: An analysis of price-only contracts. *Manufacturing & Service Operations Management* 3, No.4, 293–305.

[14] Matsuyama K., 2004. The multi-period newsboy problem. *European Journal of Operational Research* 171, 1, 170–188.

[15] Mills, E. S., 1959. Uncertainty and price theory. *Quarterly Journal of Economics* 73, 116-130.

[16] Øksendal B. and A. Sulem, 2007. *Applied Stochastic Control of Jump Diffusions.* Second Edition. Springer, Berlin Heidelberg New York.

[17] Øksendal B. and A. Sulem, 2012. Forward-backward SDE games and stochastic control under model uncertainty. *Journal of Optimization Theory and Applications.* DOI 10.1007/s10957-012-0166-7.

[18] Pardoux E. and S. Peng, 1990. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters* 14, 55–61.

[19] Perakis G. and G. Roels, 2008. Regret in the newsvendor model with partial information. *Operations Research* 56, 1, 188–203.

[20] Petruzzi, C. N. and M. Dada, 1999. Pricing and the newsvendor problem: A review with extensions. *Operations Research* 47, 183–194.

[21] Qin Q., R. Wang, A. J. Vakharia, Y. Chen, and M. Seref, 2011. The newsvendor problem: Review and directions for future research. *European Journal of Operational Research* 213, 361–374.

[22] Scarf H., 1958. A min-max solution of an inventory problem. In: *Studies in The Mathematical Theory of Inventory and Production*, K. Arrow, S. Karlin, and H. Scarf, (eds), Stanford: Stanford University Press, pp. 201–209.

[23] Wang H., B. Chen, and H. Yan, 2010. Optimal inventory decisions in a multiperiod newsvendor model with partially observed Markovian supply capacities. *European Journal of Operational Research* 202, No. 2, 502–517.

[24] Whitin, T.M., 1955. Inventory control and price theory. *Management Science* 2, 61–68.