Discussion paper

Stochastic Stackelberg equilibria with applications to time dependent newsvendor models

BY
BERNT ØKSENDAL, LEIF SANDAL, AND **JAN UBØE**

Stochastic Stackelberg equilibria with applications to time dependent newsvendor models

Bernt Øksendal*, Leif Sandal and Jan Ubøe

May 18, 2011

Abstract

In this paper we prove a sufficient 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. We provide complete existence and uniqueness proofs for a series of special cases, including geometric Brownian motion and the Ornstein-Uhlenbeck process, both with time variable coefficients. Moreover, these results are operational because we are able to offer explicit solution formulas. An interesting finding is that more precise information may be a considerable disadvantage for the retailer.

Keywords:

Stochastic differential games, newsvendor model, delayed information, Itô-Lévy processes

Bernt Øksendal: Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, 0316 Oslo, Norway, e-mail: oksendal@math.uio.no

Leif Sandal: Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway, e-mail: leif.sandal@nhh.no

Jan Ubøe: Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway, e-mail: jan.uboe@nhh.no

Main variables:

$$w = \text{ wholesale price per unit (chosen by the manufacturer)}$$
$$q = \text{ order quantity (rate chosen by the retailer)}$$
$$R = \text{ retail price per unit (chosen by the retailer)}$$
$$D = \text{ demand (random rate)}$$
$$M = \text{ production cost per unit (fixed)}$$
$$S = \text{ salvage price per unit (fixed)}$$

1 Introduction

The one periodic newsvendor model is a widely studied object that has attracted increasing interest in the two last decades. The basic setting is that a retailer wants to order a quantity q from a manufacturer. The demand D is a random variable, and the retailer wishes to select an order quantity q maximizing his expected profit $\mathrm{E}[\Pi_r(q, D)]$. 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 by now a huge literature on such problems, and for further reading we refer to the survey papers by Cachon (2003) and Qin et. al (2011) and the numerous references therein.

The (discrete) multiperiod newsvendor problem has been studied in detail by many authors, Matsuyama (2004), Berling (2006), Bensoussan et. al (2007, 2009), Wang et. al (2010), just to quote some of the more recent contributions. Two papers that come somewhat close to the approach used in our paper are Kogan (2003) and Kogan and Lou (2003) , where the authors consider continuous time scheduling problems.

In many cases the 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, how-

2

ever, 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 the 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 the size of a delivery rate. The manufacturer is the leader and decides the wholesale price. Based on the given wholesale price, the retailer decides the size of the delivery rate. 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 obviously the retailer puts his order rate 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.

Multiperiod newsvendor problems with delayed information have been discussed in several papers, but none of these papers appears to make the theory operational. Bensoussan el. al (2009) use a time-discrete approach and generalizes several information delay models, however, under the assumption of independence of the delay process from the inventory, the demand, and the ordering process. They write 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 leads to non-trivial numerical problems 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.

Figure 1 shows a sample path of an Ornstein-Uhlenbeck process which 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.
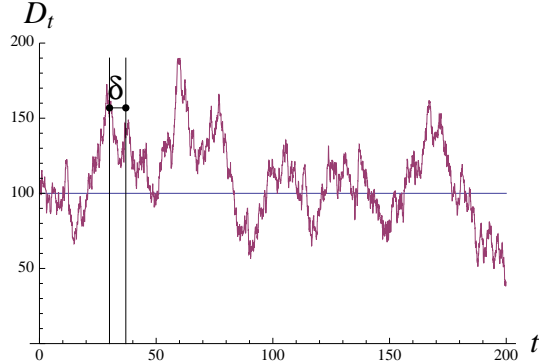

Figure 1: An Ornstein-Uhlenbeck process with delayed information

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. Such cases are hard to solve in terms of explicit expressions, however, and 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 the 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 previously been discussed in the literature. Before we can discuss game equilibria for the newsvendor problem, we must hence formulate and prove a sufficient maximum principle for general stochastic differential Stackelberg games.

The paper is organized as follows. In Section 2 we set up a framework where we discuss general stochastic differential Stackelberg games. To keep the discussion simple, we only discuss the problem in very broad terms. The details are technically demanding, and the complete proof and formulation of our maximum principle is placed in Appendix A. In Section 3 we discuss a continuous time newsvendor problem. As the newsvendor problem is a special case of the general framework in Section 2, the results in Appendix A can be used to formulate explicit strategies for finding equilibria of such problems. In Section 4 we consider the case where the demand rate is given by an Ornstein-Uhlenbeck process, and are able to provide explicit solutions for the unique equilibria that occurs in that case. Section 5 is devoted to geometric Brownian motion, and explicit solutions are provided for that case as well. The result in the constant coefficient case is quite startling, as it leads to an equilibrium where the manufacturer offers a constant price w and the retailer orders a fixed fraction of the observed demand rate. In Section 6 we compare the dynamic approach with a static approach where both parties (wrongly) believes that the demand rate has a static distribution. An interesting finding is that more precise information can be a considerable disadvantage for the retailer. Finally in Section 7 we offer some concluding remarks.

As already mentioned above, parts of the paper are technically demanding. To make the paper available to a larger public, complete proofs are in most cases placed in the Appendices. Appendix A provides the full details for stochastic differential Stackelberg games, while Appendix B contains all the rest of the proofs.

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 can at time t choose a control $u_1(t)$ while player 2 can choose a control $u_2(t)$. The controls determine how X_t evolves in time. We will use bolded characters to denote functions, e.g., $\mathbf{u}_1$ denotes the function $u_1(t), 0 \leq t \leq T$. The performance for player i is assumed to be on the form

$$J_i(\mathbf{u}_1, \mathbf{u}_2) = \mathrm{E}\left[\int_0^T f_i(t, X_t, u_1(t), u_2(t))dt + g_i(X_T)\right] \qquad i = 1, 2 \qquad (1)$$

where f_1, f_2, g_1, g_2 are given functions.

In our Stackelberg game player 1 is the leader, and player 2 the follower. Hence when $\mathbf{u}_1$ is revealed to player 2, player 2 will choose $\mathbf{u}_2$ to maximize $J_2(\mathbf{u}_1, \mathbf{u}_2)$. Player 1 knows that player 2 will act in this rational way.

Suppose that for any given control $\mathbf{u}_1$ there exists a map Φ that selects $\mathbf{u}_2$ which maximizes $J_2(\mathbf{u}_1, \mathbf{u}_2)$. Player 1 will hence choose $\mathbf{u}_1 = \mathbf{u}_1^*$ such that $\mathbf{u}_1 \mapsto J_1(\mathbf{u}_1, \Phi(\mathbf{u}_1))$ is maximal for $\mathbf{u}_1 = \mathbf{u}_1^*$. In order to solve problems of this type we need to specify how the state of the system evolves in time. We will work in a framework where X_t is a controlled jump diffusion on the form

$$
\begin{aligned}
dX_t = {}& \mu(t, X_t, u_1(t), u_2(t), \omega)dt + \sigma(t, X_t, u_1(t), u_2(t), \omega)dB_t \\
& + \int_{\mathbb{R}} \gamma(t, X_t, u_1(t), u_2(t), \xi, \omega)\tilde{N}(dt, d\xi) \\
X(0) = {}& x \in \mathbb{R}
\end{aligned}
\tag{2}
$$

where μ, σ, γ are given continuous functions, assumed to be continuously differentiable with respect to the variables X, u_1 and u_2. More precise definitions are provided in Appendix A. See also Øksendal and Sulem (2007) for more information about Lévy processes and controlled jump diffusions. The framework defined by (2) is very general, and contains among many other things, important special cases like the Ornstein-Uhlenbeck process and geometric Brownian motion.

To solve optimization problems related to processes on the form given in (2), one needs to formulate Hamiltonians and write down and solve the adjoint equations. Anticipating that many readers are unfamiliar with such optimization methods, these technical parts are discussed in full detail in Appendix A. At this point it suffices to say that a sufficient maximum principle can be found. In Section 3 we will consider the particular case of the newsvendor model. This case is more explicit, and the machinery in Appendix A can be used to make a set of explicit equations that can be applied without reference to the underlying details in Appendix A. It is hence possible to understand and analyze large parts of this paper without venturing the depths of Appendix A.

3 A continuous time newsvendor problem

In this section we will formulate a continuous time newsvendor problem and use the results in Appendix A 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 production 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. The assumption that items cannot be stored is of course quite restrictive. There are still many important cases leading to an 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

$$J_2(\mathbf{w}, \mathbf{q}, \mathbf{R}) = \mathrm{E}\left[\int_\delta^T (R_t - S)\min[D_t, q_t] - (w_t - S)q_t dt\right] \tag{3}$$

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

$$J_1(\mathbf{w}, \mathbf{q}, \mathbf{R}) = \mathrm{E}\left[\int_\delta^T (w_t - M)q_t dt\right] \tag{4}$$

Technical remark

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 ago. Similarly, when the 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 of course 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) jump diffusion on the form:

$$dD_t = \mu(t, D_t, R_t)dt + \sigma(t, D_t, R_t)dB_t + \int_{\mathbb{R}} \gamma(t, D_t, R_t, \xi)\tilde{N}(dt, d\xi); \qquad t \in [0, T] \qquad (5)$$

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

Brownian motion B_t and the compensated Poisson term $\tilde{N}(t, dz)$ are driving the stochastic differential equation in (5), and it is hence natural to formalize information with respect to these objects. We hence 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

$$\mathcal{E}_t := \mathcal{F}_{t-\delta} \qquad t \in [\delta, T] \qquad (6)$$

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., $\mathbf{q}$ and $\mathbf{w}$ should be $\mathcal{E}$-predictable processes. In less technical terms $\mathcal{E}$-predictable means that for each t it must be possible to write down the value based on a limit of functions of the values of the Brownian motion and the compensated Poisson term at times up to $t - \delta$. As mentioned above, the retail price R_t can be settled at a later stage. It is hence possible to consider a second delay $\delta_R < \delta$, and assume that R_t is decided at time $t - \delta_R$. We hence consider a second σ-algebra $\mathcal{E}_t^R := \mathcal{F}_{t-\delta_R}$, $t \in [\delta, T]$ and will assume that $\mathbf{R}$ is $\mathcal{E}^R$-predictable.

3.2 Finding Stackelberg equilibria in the newsvendor model

Using the machinery developed in Appendix A, we can now write down a system of equations defining Stackelberg equilibria for our newsvendor model. We will focus on the simplified case where D_t does not depend on the retail price, and where $R_t = R$ is exogeneously given and fixed. In that particular case the resulting equations are quite transparent and the following

8

theorem summarizes the result.

Theorem 3.2.1

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

$$\mathrm{E}\left[(R - S)\mathcal{X}_{[0,D_t]}(q_t) - w_t + S|\mathcal{E}_t\right] = 0 \tag{7}$$

Let $q_t = \phi(w_t)$ denote the unique solution of (7), and assume that the function

$$w_t \mapsto \mathrm{E}\left[(w_t - M)\phi(w_t)|\mathcal{E}_t\right] \tag{8}$$

has unique maximum at $w_t = \hat{w}_t$. If $\hat{q}_t = \phi(\hat{w}_t)$, then the pair $(\hat{\mathbf{q}}, \hat{\mathbf{w}})$ is a Stackelberg equilibrium for the newsvendor problem defined by (4) and (3).

Here $\mathcal{X}_{[0,D_t]}(q)$ denotes the indicator function for the interval $[0, D_t]$, i.e., a function that has the value 1 if $0 \le q \le D_t$, and is zero otherwise. To see why (7) always has a unique solution, note that w_t is $\mathcal{E}_t$-measurable and hence (7) is equivalent to

$$\mathrm{E}\left[\mathcal{X}_{[0,D_t]}(q_t)|\mathcal{E}_t\right] = \frac{w_t - S}{R - S} \tag{9}$$

Existence and uniqueness of q_t then follows from monotonicity of conditional expectation. 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 (7) and prove that (8) has unique maxima. Notice that (7) 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.

In Appendix A we offer a complete solution to the general case where D_t is a function of R_t, and R_t is a decision variable. In the general case the equilibria satisfy a set of non-linear stochastic differential equations. Theorem 3.2.1 will be sufficient for many of the applications we have in mind, however.

4 Explicit formulas for the Ornstein-Uhlenbeck process

In this section we offer explicit formulas for the equilibria that occur when the demand rate is given by an Ornstein-Uhlenbeck process. We first consider the case with constant coefficients, and then extend the results to the case with time dependent, deterministic coefficients.

4.1 Ornstein-Uhlenbeck process with constant coefficients

In this section we assume that D_t is an Ornstein-Uhlenbeck process with constant coefficients, i.e., that

$$dD_t = a(\mu - D_t)dt + \sigma dB_t \tag{10}$$

where a, μ, 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 (10) is

$$D_t = D_0 e^{-at} + \mu(1 - e^{-at}) + \int_0^t \sigma e^{a(s-t)} dB_s \tag{11}$$

It is easy to see that

$$D_t = D_{t-\delta} e^{-a\delta} + \mu(1 - e^{-a\delta}) + \int_{t-\delta}^t \sigma e^{a(s-t)} dB_s \tag{12}$$

Using that 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 possible to find a closed form solution to (7). 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

$$\Phi_y[w] = ye^{-a\delta} + \mu(1 - e^{-a\delta}) + \sigma\sqrt{\frac{1 - e^{-2a\delta}}{2a}} \cdot G^{-1}\left[1 - \frac{w - S}{R - S}\right] \tag{13}$$

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 quasi-concave and has a unique maximum with a strictly positive function value.

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

$$w_t^* = \begin{cases} \text{Argmax}[\Psi_y] & \text{if } \Phi_y[M] > 0 \\ M & \text{otherwise} \end{cases} \qquad q_t^* = \begin{cases} \Phi_y[\text{Argmax}[\Psi_y]] & \text{if } \Phi_y[M] > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (14)$$

Proof

See Appendix B. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

The condition $\Phi_y[M] > 0$ has an obvious interpretation. The manufacturer cannot offer a wholesale price w lower than her 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 retailers best strategy is to order $q = 0$ units. When the retailer orders $q = 0$ units, the choice of w is arbitrary. The choice $w = M$ is, however, 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 (13) 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

$$a = 0.05 \qquad \mu = 100 \qquad \sigma = 12 \qquad R = 10 \qquad S = 1 \qquad M = 2 \qquad (15)$$

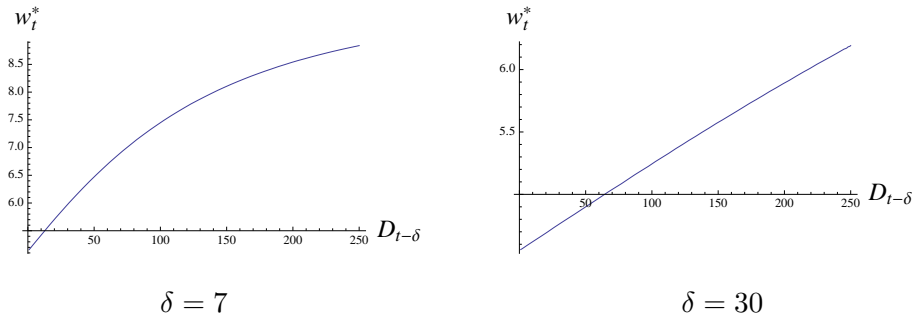$$\delta = 7 \qquad\qquad\qquad\qquad\qquad\qquad\qquad \delta = 30$$

Figure 2: w_t^* as a function of the observed demand rate $D = D_{t-\delta}$
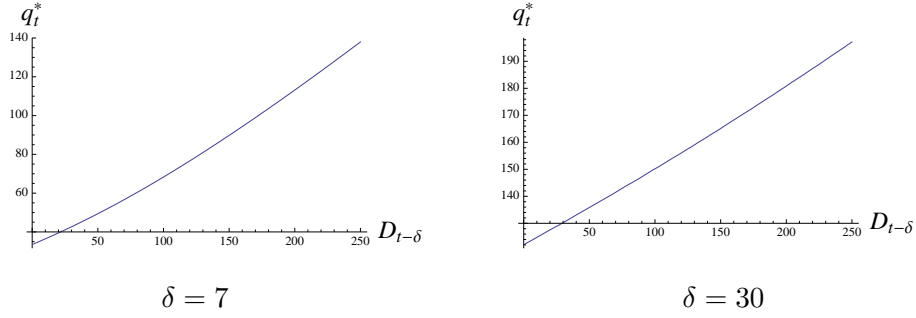
Figure 3: q_t^* as a function of the observed demand rate $D = D_{t-\delta}$

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 Ornstein-Uhlenbeck process with variable coefficients

We now consider the Ornstein-Uhlenbeck process with variable, deterministic coefficients, i.e., the case

$$dD_t = a(t)(\mu(t) - D_t)dt + \sigma(t)dB_t \tag{16}$$

where $a(t), \mu(t), \sigma(t)$ are given deterministic functions. The increased flexibility is important in applications since it allows for scenarios where the mean reversion level μ can have a time variable trend. When the coefficients are constant, the equilibria can be found by simple look-up tables as the ones shown in Figure 2 and 3. Moreover, the look-up tables are the same for all values of t. This is no longer true in the case with variable coefficients. The equilibria can still be found from look-up tables, but these tables are in general different for different values of t. The basic result can be summarized as follows:

Proposition 4.2.1

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

$$\Phi_{t,y}[w] = ye^{-\int_{t-\delta}^{t} a(s)ds} + \hat{\mu}(t) + \hat{\sigma}(t)G^{-1}\left[1 - \frac{w - S}{R - S}\right] \tag{17}$$

where

$$\hat{\mu}(t) = \int_{t-\delta}^{t} a(s)\mu(s)e^{-\int_s^t a(u)du}ds \qquad \hat{\sigma}(t) = \sqrt{\int_{t-\delta}^{t} \sigma^2(s)e^{-2\int_s^t a(u)du}ds} \tag{18}$$

12

and let $\Psi_{t,y} : [M, R] \to \mathbb{R}$ denote the function $\Psi_{t,y}[w] = (w - M)\Phi_{t,y}[w]$. If $\Phi_{t,y}[M] > 0$, the function $\Psi_{t,y}$ is quasi-convave and has a unique maximum with a strictly positive function value.

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

$$
w_t^* = \begin{cases} \mathrm{Argmax}[\Psi_{t,y}] & \textit{if } \Phi_{t,y}[M] > 0 \\ M & \textit{otherwise} \end{cases} \qquad q_t^* = \begin{cases} \Phi_{t,y}[\mathrm{Argmax}[\Psi_{t,y}]] & \textit{if } \Phi_{t,y}[M] > 0 \\ 0 & \textit{otherwise} \end{cases}
$$

$$(19)$$

Proof

See Appendix B. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

As in the constant coefficient case, these formulas are sufficiently explicit to make the model operational. The unique values $\mathrm{Argmax}[\Psi_{t,y}]$ are easily found numerically, and look-up tables similar to the ones shown in Figure 2 and 3 can be constructed for each fixed t.

5 Explicit formulas for geometric Brownian motion

In this section we offer explicit formulas for the equilibria that occur when the demand rate is given by a geometric Brownian motion. We first consider the case with constant coefficients, and then extend the results to the case with time dependent, deterministic coefficients.

While geometric Brownian motion is probably the most commonly used model for stock prices, it is not necessarily an obvious choice to model a demand rate. The Ornstein-Uhlenbeck process is often the preferred choice allowing demand to fluctuate around a pre-determined mean-reversion level. Ornstein-Uhlenbeck processes can, however, take on negative values with positive probability, and that problem never occurs with geometric Brownian motion. If we model demand by a geometric Brownian motion, we are basically assuming that demand is growing proportional to the size of the market. That assumption may be unreasonable in a settled market, but can be attractive in modeling emerging markets with newly introduced brands.

5.1 Geometric Brownian motion with constant coefficients

In this section we assume that D_t is a geometric Brownian motion with constant coefficients, i.e., that

$$dD_t = aD_t dt + \sigma D_t dB_t \tag{20}$$

where a, σ are constants. The explicit solution to (20) is

$$D_t = D_0 \exp\left[\left(a - \frac{1}{2}\sigma^2\right)t + \sigma B_t\right] \tag{21}$$

and it is easy to see that

$$D_t = D_{t-\delta} \exp\left[\left(a - \frac{1}{2}\sigma^2\right)\delta + \sigma(B_t - B_{t-\delta})\right] \tag{22}$$

The explicit form of (22) makes it possible to write down a closed form solution to (7):

Proposition 5.1.1

Let $\Phi : [M, R] \to \mathbb{R}$ denote the function

$$\Phi[w] = \exp\left[(a - \frac{1}{2}\sigma^2)\delta + \sqrt{\delta\sigma^2} \cdot G^{-1}\left[1 - \frac{w - S}{R - S}\right]\right] \tag{23}$$

and let $\Psi : [M, R] \to \mathbb{R}$ denote the function $\Psi[w] = (w - M)\Phi[w]$. The function Ψ is quasi-concave and has a unique maximum with a strictly positive function value.

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

$$w_t^* = \mathrm{Argmax}[\Psi] \quad (constant) \qquad\qquad q_t^* = y \cdot \Phi[\mathrm{Argmax}[\Psi]] \tag{24}$$

Proof

See Appendix B. □

We see that the solution is quite different from the case with the Ornstein-Uhlenbeck process.

In this case the manufacturer has an equilibrium price that is constant and need not observe the demand rate to decide her price. The retailer should order a fixed fraction of the observed demand rate at time $t - \delta$.

5.2 Geometric Brownian motion with variable coefficients

In this section we assume that D_t is a geometric Brownian motion with variable deterministic coefficients, i.e., that

$$dD_t = a(t)D_t dt + \sigma(t)D_t dB_t \tag{25}$$

where $a(t), \sigma(t)$ are given deterministic functions. The increased flexibility is of course attractive in modeling, but like in the Ornstein-Uhlenbeck case, the solution is more complicated.

Proposition 5.2.1

For $t \in [\delta, T]$, let $\Phi_t : [M, R] \to \mathbb{R}$ denote the function

$$\Phi_t[w] = \exp\left[\hat{a}(t) + \hat{\sigma}(t) \cdot G^{-1}\left[1 - \frac{w - S}{R - S}\right]\right] \tag{26}$$

where

$$\hat{a}(t) = \int_{t-\delta}^{t} a(s) - \frac{1}{2}\sigma^2(s)ds \qquad \hat{\sigma}(s) = \sqrt{\int_{t-\delta}^{t} \sigma^2(s)dt} \tag{27}$$

and let $\Psi_t : [M, R] \to \mathbb{R}$ denote the function $\Psi_t[w] = (w - M)\Phi_t[w]$. The function Ψ_t is quasi-concave and has a unique maximum with a strictly positive function value.

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

$$w_t^* = \text{Argmax}[\Psi_t] \qquad\qquad q_t^* = y \cdot \Phi_t[\text{Argmax}[\Psi_t]] \tag{28}$$

Proof

See Appendix B. □

If we compare with the case with constant coefficients, we see that the wholesale price w is no

longer constant. Nevertheless we see that the equilibria are defined in terms of two deterministic functions $\text{Argmax}[\Psi_t]$ and $\Phi_t[\text{Argmax}[\Psi_t]]$. The values for these functions are easily computed numerically. To illustrate how this works, we used the parameter values

$$R = 10, S = 1, M = 2, \delta = 7$$

together with the coefficients

$$a(t) = 0.02 \text{ (constant)} \qquad \sigma(t) = 0.075 \cdot (e^{-0.05\,t} + 1) \tag{29}$$

Graphs for the functions $\text{Argmax}[\Psi_t]$ and $\Phi_t[\text{Argmax}[\Psi_t]]$ are shown in Figure 4.

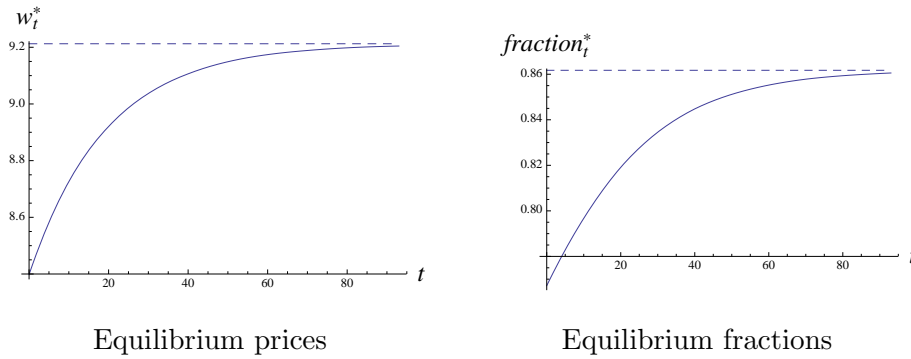Equilibrium prices   Equilibrium fractions

Figure 4: Equilibrium prices and fractions for geometric Brownian motion

The dotted line shows the corresponding equilibrium price and fraction for the constant coefficient case with

$$a = 0.02 \qquad \sigma = 0.075 \tag{30}$$

5.3   Geometric Lévy processes with ω dependent coefficients

In this subsection we handle geometric Lévy processes where one of the coefficients is ω dependent. This case is important since it focuses some general problems, and also provide some hints on how to address even more general cases. A geometric Lévy process is a solution of a stochastic differential equation of the form

$$dD(t) = D(t^-) \left(a(t,\omega)dt + \sigma(t,\omega)dB_t + \int_{\mathbb{R}} \gamma(t,z,\omega)\tilde{N}(dt,dz) \right) \tag{31}$$

In this equation B_t is Brownian motion and $\tilde{N}$ denotes a compensated Poisson random mea-

sure, see Øksendal and Sulem (2007) for an introduction to stochastic calculus based on Lévy processes. If we assume that $D(0) = D_0 > 0$ and $\gamma(t, z) \geq -1$, the solution satisfies $D_t \geq 0$ for all t. The model in (31) is widely accepted as a good model for prices of financial assets. The positivity property $D_t \geq 0$ for all t, makes it suitable as a model for demand as well. If $\gamma(t, z, \omega) \in [-1, 0]$ for all (t, z, ω), this corresponds to a situation with market corrections, i.e., at random times demand is adjusted to a fraction of its previous value. This is an attractive way to negate the exponential growth of geometric Brownian motion, keeping demand from growing without limit.

The explicit solution of (31) is

$$D_t = D_0 \exp \left[\int_0^t \left(a(s, \omega) - \frac{1}{2}\sigma^2(s, \omega) + \int_{\mathbb{R}_0} \log[1 + \gamma(s, z, \omega)] - \gamma(s, z, \omega)\nu(dz) \right) ds \right. \quad (32)$$
$$\left. + \int_0^t \sigma(s, \omega)dB_s + \int_0^t \int_{\mathbb{R}_0} \log[1 + \gamma(s, z, \omega)]\tilde{N}(ds, dz) \right]$$

Now assume that

$$a(s, \omega) = B_s(\omega) \qquad \sigma(s, \omega) = \sigma(s) \qquad \gamma(s, z, \omega) = \gamma(s, z) \qquad (33)$$

i.e., that σ and γ are given deterministic functions, while the growth-rate $a(s, \omega)$ is depending om ω as well as t. The expression in (32) is sufficiently explicit to admit a fair description of the functions we need to solve to find Stackelberg equilibria. This description can be formulated as follows:

For each fixed t, consider the random variable X_t given by

$$X_t = \exp \left[\delta B_{t-\delta} + t(B_t - B_{t-\delta}) - \int_{t-\delta}^t s dB_s \right.$$
$$+ \int_{t-\delta}^t \left(-\frac{1}{2}\sigma^2(s) + \int_{\mathbb{R}_0} \log[1 + \gamma(s, z)] - \gamma(s, z)\nu(dz) \right) ds \quad (34)$$
$$\left. + \int_{t-\delta}^t \sigma(s)dB_s + \int_{t-\delta}^t \int_{\mathbb{R}_0} \log[1 + \gamma(s, z)]\tilde{N}(ds, dz) \right]$$

Let F_t denote the cumulative distribution of X_t, and for each fixed t let F_t^{-1} denote the inverse

function of F_t. Consider for each $t \in [\delta, T]$, the functions

$$\Phi_t[w] = F_t^{-1} \left[1 - \frac{w - S}{R - S} \right] \qquad \Psi_t[w] = (w - M)\Phi_t[w] \qquad (35)$$

At time $t - \delta$ the retailer should observe both the demand rate $y = D_{t-\delta}$ and the growth rate $a = B_{t-\delta}$, and equilibrium is obtained at

$$w_t^* = \text{Argmax}[\Psi_t] \qquad q_t^* = y \cdot e^{\delta a} \cdot \Phi_t[\text{Argmax}[\Psi_t]] \qquad (36)$$

For a derivation of these results see Appendix B. We notice that the structure of the solution is quite similar to the case covered in Proposition 5.2.1. The manufacturer has a pricing strategy defined in terms of a deterministic function. The retailer should observe the demand rate, adjust it by the observed growth rate, and multiply the adjusted number by a deterministic fraction.

A problem here is that the random variable in (34) is terribly complicated leaving little hope of an analytic proof of a unique maximum. Apart from that, the expressions can still be handled numerically, and maxima can be verified by visual inspection.

6 Numerical examples for the Ornstein-Uhlenbeck process

In this section we will compare the performance of the dynamic approach with a scenario where the retailer believes that demand has a constant distribution D. A constant coefficient Ornstein-Uhlenbeck process

$$D_t = D_0 e^{-at} + \mu(1 - e^{-at}) + \sigma e^{-at} \cdot \int_0^t e^{as} dB_s \qquad (37)$$

is ergodic in the sense that observations along any sample path will approach the distribution $N(\mu, \frac{\sigma^2}{2a})$. Assuming that the retailer believes that the demand rate has a static distribution D and that he has observed the demand rate a fair amount of time prior to ordering, he will hence conclude that D is $N(\mu, \frac{\sigma^2}{2a})$. If the manufacturer knows that the retailer will make orders according to a static $N(\mu, \frac{\sigma^2}{2a})$ distribution, she can compute a fixed value for w which optimizes her expected profit.

To examine the performance of the dynamic and the static approch, we sampled paths of

18

the Ornstein-Uhlenbeck process using the parameters

$$\mu = 100 \qquad \sigma = 12 \qquad a = 0.05 \qquad D_0 = 100 \tag{38}$$

One such sample path was shown in Figure 1. Values for the accumulated profits

$$\int_\delta^T (R - S) \min[D_t, q_t] - (w_t - S)q_t dt \qquad \qquad \int_\delta^T (w_t - M)q_t dt \tag{39}$$

were computed for different values of δ using the values

$$R = 10 \qquad M = 2 \qquad S = 1 \qquad T = 100 + \delta \tag{40}$$

and using 4 different strategies:

- Dynamic approach as defined by Proposition 4.1.1.

- Static approach as defined above.

- Dynamic cooperative approach using $w_t = M$.

- Static cooperative approach using $w_t = M$.

We assume that sales takes place in time intervals $[\delta, 100 + \delta]$. The length of the sales period hence is 100 regardless of the value on δ. This makes it easier to compare performance using different values of δ. The results were averaged over 1000 sample paths and these averages are reported in the tables below.

Table 1: Performance of dynamic versus static strategies. Delay $\delta = 1$.			
Values over 1000 sample paths	Manufacturer	Retailer	Supply chain
Average profit static approach	42 830	12 729	55 559
Average profit dynamic approach	61 356	4 073	65 429
Average profit static cooporation	-	-	73 251
Average profit dynamic cooporation	-	-	77 766

Table 2: Performance of dynamic versus static strategies. Delay $\delta = 7$.			
Values over 1000 sample paths	Manufacturer	Retailer	Supply chain
Average profit static approach	42 830	12 457	55 286
Average profit dynamic approach	48 592	9 438	58 030
Average profit static cooporation	-	-	73 029
Average profit dynamic cooporation	-	-	74 838

Table 3: Performance of dynamic versus static strategies. Delay $\delta = 30$.			
Values over 1000 sample paths	Manufacturer	Retailer	Supply chain
Average profit static approach	42 830	12 074	54 903
Average profit dynamic approach	43 225	11 882	55 106
Average profit static cooporation	-	-	72 648
Average profit dynamic cooporation	-	-	72 794

As we can see from these tables, the dynamic approach is favorable for the manufacturer, and more favorable the shorter the delay. At $\delta = 30$ the effect of the dynamic approach is close to being wiped out. The same results apply for the supply chain, i.e., a dynamic approach offers improved profits and the improvement is bigger when the delay is shorter. It is interesting to note, however, that the retailer has a distinct disadvantage of the dynamic approach and that this disadvantage is bigger the shorter the delay.

In a cooperative setting a dynamic approach can be favorable for both the retailer and the manufacturer. Profits can be shared leading to an improved position for both parties. In a non-cooperative equilibrium more precise information can be a disadvantage for the retailer.

This is due to the Stackelberg structure of the game. With more precise information, the leader is in more control and can take a bigger share of the profits. In the limit $\delta \to 0$, the leader is in full control. The retailer will then order the observed demand rate no matter what the price is. The manufacturer offers a price marginally close to R taking all profit in the limit.

7   Concluding remarks

This paper has two main topics. First we develop a new theory for stochastic differential Stackelberg games and apply that theory to continuous time newsvendor problems. Second we make our theory operational by providing explicit solution formulas for important special cases.

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 A.2 in the appendix. The solution is defined in terms of a coupled system of stochastic differential equations, however, and admittedly such systems are very hard to solve.

To be able to solve our problem in terms of explicit expressions, we need to consider the simplified case where R_t is exogenously given and fixed. We offer closed form solutions for the Ornstein-Uhlenbeck process and Geometric Brownian motion both with time-variable, deterministic coefficients. We also offer formulas for a Geometric Lévy process, i.e., a process with jumps. Although we do not state that result explicitly, the discussion in Subsection 5.3 really shows how to solve all cases where the Geometric Lévy process has time-variable, deterministic coefficients.

From an applied point of view we believe that the numerical results in Section 6 are of general interest. We demonstrate that the retailer suffers a distinct disadvantage of more information, and that this disadvantage is bigger the more precise the information is. Such issues may have important political implications, in particular in electricity markets, and we believe that our model may offer new insights into the mechanisms governing equilibria in such markets.

8   Appendix A

In this appendix we offer a full discussion and complete proofs for general stochastic Stackelberg games. To improve readability we repeat some definitions given in the main body of the paper. We will assume that the state of the system is given by a controlled jump diffusion on the form:

$$
dX_t = \mu(t, X_t, u(t), \omega)dt + \sigma(t, X_t, u(t), \omega)dB_t
$$

$$
+ \int_{\mathbb{R}} \gamma(t, X_t, u(t), \xi, \omega)\tilde{N}(dt, d\xi) \tag{41}
$$

$$
X(0) = x \in \mathbb{R}
$$

where the coefficients $\mu(t, x, u, \omega) : [0, T] \times \mathbb{R} \times \mathbb{U} \times \Omega \to \mathbb{R}$, $\sigma(t, x, u, \omega) : [0, T] \times \mathbb{R} \times \mathbb{U} \times \Omega \to \mathbb{R}$, $\gamma(t, x, u, \xi, \omega) : [0, T] \times \mathbb{R} \times \mathbb{U} \times \mathbb{R}_0 \times \Omega \to \mathbb{R}$ are given continuous functions assumed to be continuously differentiable with respect to x and u, and $\mathbb{R}_0 = \mathbb{R} \backslash \{0\}$. Here $B_t = B(t, \omega); (t, \omega) \in [0, \infty) \times \Omega$ is a Brownian motion 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 (2005) for more information about controlled jump diffusions. The set $\mathbb{U} = \mathbb{U}_1 \times \mathbb{U}_2$ is a given set of admissible control values $u(t, \omega)$. We assume that the control $u = u(t, \omega)$ consists of two components, $u = (u_1, u_2)$, where Player 1 controls u_1 and Player 2 controls u_2. We also assume that the information flows available to the players are given filtrations $\{\mathcal{E}_t^{(1)}\}_{t \in [0,T]}$ for control u_1 and control $u_2^{(1)}$ and $\{\mathcal{E}_t^{(2)}\}_{t \in [0,T]}$ for control $u_2^{(2)}$, where $u_2 = (u_2^{(1)}, u_2^{(2)})$ and

$$
\mathcal{E}_t^{(1)} \subseteq \mathcal{E}_t^{(2)} \subseteq \mathcal{F}_t \qquad \text{for all } t \in [0, T]. \tag{42}
$$

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

$$
\mathcal{E}_t^{(1)} = \mathcal{E}_t^{(2)} = \mathcal{F}_{t-\delta} \qquad \text{for all } t \in [\delta, T]. \tag{43}
$$

for some fixed information delay $\delta > 0$. We assume that $u_1(t)$ and $u_2^{(1)}(t)$ are $\mathcal{E}_t^{(1)}$-predictable, and that $u_2^{(2)}(t)$ is $\mathcal{E}_t^{(2)}$-predictable. Hence we assume there are three given families $\mathcal{A}_{\mathcal{E}}^{(1)}, \mathcal{A}_{\mathcal{E}}^{(2,1)}$ and $\mathcal{A}_{\mathcal{E}}^{(2,2)}$ of admissable controls $u_1, u_2^{(1)}, u_2^{(2)}$, contained in the set of $\mathcal{E}_t^{(1)}, \mathcal{E}_t^{(1)}$ and $\mathcal{E}_t^{(2)}$-predictable processes, respectively, and we put

$$\mathcal{A}_{\mathcal{E}} = \mathcal{A}_{\mathcal{E}}^{(1)} \times \mathcal{A}_{\mathcal{E}}^{(2)} \qquad \text{where } \mathcal{A}_{\mathcal{E}}^{(2)} = \mathcal{A}_{\mathcal{E}}^{(2,1)} \times \mathcal{A}_{\mathcal{E}}^{(2,2)} \tag{44}$$

If a control $u = (u_1, u_2) \in \mathcal{A}_{\mathcal{E}}$ is chosen, the performance for Player i is assumed to be of the form

$$J_i(u) = \mathrm{E}\left[\int_0^T f_i(t, X_t, u(t))dt + g_i(X_T)\right] \qquad i = 1, 2 \tag{45}$$

where $f_i(t, x, u) : [0, T] \times \mathbb{R} \times \mathbb{U} \to \mathbb{R}$ and $g_i(x) : \mathbb{R} \to \mathbb{R}$ are given profit rates and bequest functions, respectively, assumed to be C^1 with respect to x and u.

We now consider the following game theoretic situation:

Suppose Player 1 decides her control process $u_1 \in \mathcal{A}_{\mathcal{E}}^{(1)}$. At any time t the value is immediately known to Player 2. Therefore he chooses $u_2 = u_2^* \in \mathcal{A}_{\mathcal{E}}^{(2)}$ such that

$$u_2 \mapsto J_2(u_1, u_2) \text{ is maximal for } u_2 = u_2^*. \tag{46}$$

Assume the there exists a measurable map $\Phi : \mathcal{A}_{\mathcal{E}}^{(1)} \to \mathcal{A}_{\mathcal{E}}^{(2)}$ (a "maximizer" map) such that

$$u_2 \mapsto J_2(u_1, u_2) \text{ is maximal for } u_2 = u_2^* = \Phi(u_1) \tag{47}$$

Player 1 knows that Player 2 will act in this rational way. Therefore Player 1 will choose $u_1 = u_1^* \in \mathcal{A}_{\mathcal{E}}^{(1)}$ such that

$$u_1 \mapsto J_1(u_1, \Phi(u_1)) \text{ is maximal for } u_1 = u_1^*. \tag{48}$$

The control $u^* := (u_1^*, \Phi(u_1^*)) \in \mathcal{A}_{\mathcal{E}}^{(1)} \times \mathcal{A}_{\mathcal{E}}^{(2)}$ is called a Stackelberg equilibrium for the game defined by (41)-(45). In the newsvendor problem studied in this paper Player 1 is the manufacturer who decides the wholesale price $u_1 = w$ for the retailer, who is Player 2, and who decides the order rate $u_2^{(1)} = q$ and the retailer price $u_2^{(2)} = R$. Thus $u_2 = (q, R)$. We may summarize (46) and (48) as follows:

$$\max_{u_2 \in \mathcal{A}_{\mathcal{E}}^{(2)}} J_2(u_1, u_2) = J_2(u_1, \Phi(u_1)) \tag{49}$$

and

$$\max_{u_1 \in \mathcal{A}_{\mathcal{E}}^{(1)}} J_1(u_1, \Phi(u_1)) = J_2(u_1^*, \Phi(u_1^*)) \tag{50}$$

We see that (49) and (50) constitute two consecutive stochastic control problems with partial information, and hence we can use the maximum principle for such problems (see, e.g., Framstad et. al (2004) and Baghery and Øksendal (2007)) to find a maximum principle for Stackelberg equilibria. To this end, define the Hamiltonian $H_2(t, x, u, a_2, b_2, c_2(\cdot)) : [0, T] \times \mathbb{R} \times \mathbb{U} \times \mathbb{R} \times \mathbb{R} \times \mathcal{R} \to \mathbb{R}$ by

$$H_2(t, x, u, a_2, b_2, c_2(\cdot)) = f_2(t, x, u) + \mu(t, x, u)a_2 + \sigma(t, x, u)b_2 \tag{51}$$
$$+ \int_{\mathbb{R}} \gamma(t, x, u, \xi)c_2(\xi)\nu(d\xi);$$

where $\mathcal{R}$ is the set of functions $c(\cdot) : \mathbb{R}_0 \to \mathbb{R}$ such that (51) converges. The adjoint equation for H_2 in the unknown adjoint processes $a_2(t), b_2(t)$, and $c_2(t, \xi)$ is the following backward stochastic differential equation (BSDE):

$$da_2(t) = -\frac{\partial H_2}{\partial x}(t, X(t), u(t), a_2(t), b_2(t), c_2(t, \cdot))dt \tag{52}$$
$$+ b_2(t)dB_t + \int_{\mathbb{R}} c_2(t, \xi)\tilde{N}(dt, d\xi); \quad 0 \le t \le T$$
$$a_2(T) = g_2'(X(T)) \tag{53}$$

Here $X(t) = X^u(t)$ is the solution to (2) corresponding to the control $u \in \mathcal{A}_{\mathcal{E}}$. Next, for a given map $\phi : \mathbb{R} \to \mathbb{R}$ define the Hamiltonian $H_1^\phi(t, x, u_1, a_1, b_1, c_1(\cdot)) : [0, T] \times \mathbb{R} \times \mathbb{U}_1 \times \mathbb{R} \times \mathbb{R} \times \mathcal{R} \to \mathbb{R}$ by

$$H_1^\phi(t, x, u_1, a_1, b_1, c_1(\cdot)) = f_1(t, x, u_1, \phi(u_1)) + \mu(t, x, u_1, \phi(u_1))a_1 \tag{54}$$
$$+ \sigma(t, x, u_1, \phi(u_1))b_1 + \int_{\mathbb{R}} \gamma(t, x, u_1, \phi(u_1), \xi)c_1(\xi)\nu(d\xi)$$

The adjoint equation (for H_1^ϕ) in the unknown processes $a_1(t), b_1(t), c_1(t, \xi)$ is the following BSDE:

24

$$da_1(t) = -\frac{\partial H_1^\phi}{\partial x}(t, X(t), u_1(t), \phi(u_1(t)), a_1(t), b_1(t), c_1(t, \cdot))dt \tag{55}$$

$$+ b_1(t)dB_t + \int_{\mathbb{R}} c_1(t, \xi)\tilde{N}(dt, d\xi); \quad 0 \le t \le T$$

$$a_1(T) = g_1'(X(T)) \tag{56}$$

Here $X(t) = X^{u_1, \phi(u_1)}(t)$ is the solution to (41) corresponding to the control $u(t) := (u_1(t), \phi(u_1(t)))$; $t \in [0, T]$, assuming that this is admissible.

We can now formulate our maximum principle for Stackelberg equilibria:

Theorem A.1 (Sufficient maximum principle)

i) Suppose the following bullet points hold:

- $H_2(t, x, u_1, u_2, a_2, b_2, c_2(\cdot))$ *is concave with respect to x and u_2, for each $t, u_1, a_2, b_2, c_2(\cdot)$.*

- $g_2(x)$ *is concave.*

- *There exists a function $\phi = \phi(u_1) = \phi(u_1, t, \omega) : \mathbb{U}_1 \times [0, T] \times \Omega \to \mathbb{U}_2$ such that for all $u_1 \in \mathcal{A}_{\mathcal{E}}^{(1)}$,*

$$\max_{v_2 \in \mathbb{U}_2} \mathrm{E}\left[H_2(t, \tilde{X}(t), u_1(t), v_2, \tilde{a}_2(t), \tilde{b}_2(t), \tilde{c}_2(t, \cdot))\big|\mathcal{E}_t^{(2,j)}\right] \tag{57}$$

$$= \mathrm{E}\left[H_2(t, \tilde{X}(t), u_1(t), \phi(u_1(t)), \tilde{a}_2(t), \tilde{b}_2(t), \tilde{c}_2(t, \cdot))\big|\mathcal{E}_t^{(2,j)}\right] \; ; \; j = 1, 2$$

where $\tilde{X}(t) = X^{u_1, \phi(u_1)}(t)$ and $(\tilde{a}_2, \tilde{b}_2, \tilde{c}_2)$ is the solution of (52)-(53) corresponding to $u = (u_1, \phi(u_1))$.

-

$$\mathrm{E}\left[\int_0^T (\tilde{X}(t) - X^u(t))^2 \left(\tilde{b}_2^2(t) + \int_{\mathbb{R}} \tilde{c}_2^2(t, \xi)\nu(d\xi)\right)dt \right. \tag{58}$$

$$\left. + \int_0^T \tilde{a}_2^2(t) \left(\sigma^2(t, X^u(t), u(t)) + \int_{\mathbb{R}} \gamma^2(t, X^u(t), u(t), \xi)\nu(d\xi)\right)dt\right] < \infty$$

for all $u = (u_1, u_2) \in \mathcal{A}_{\mathcal{E}}$.

Define

$$(\Phi(u_1))(t, \omega) = \phi(u_1(t, \omega)); \ u_1 \in \mathcal{A}_\mathcal{E}^{(1)} \tag{59}$$

Suppose $\Phi(u_1) \in \mathcal{A}_\mathcal{E}^{(2)}$. *Then*

$$\max_{u_2 \in \mathcal{A}_\mathcal{E}^{(2)}} J_2(u_1, u_2) = J_2(u_1, \Phi(u_1)); \ u_1 \in \mathcal{A}_\mathcal{E}^{(1)} \tag{60}$$

ii) With ϕ as in i), assume in addition that the following bullet points hold:

- $H_1^\phi(t, x, u_1, a_1, b_1, c_1(\cdot))$ *is concave with respect to x and u_1, for all $t, a_1, b_1, c_1(\cdot)$.*

- $g_1(x)$ *is concave.*

- *There exist $u_1^* \in \mathcal{A}_\mathcal{E}^{(1)}$ such that*

$$\max_{v_1 \in \mathbb{U}_1} \mathrm{E}\left[H_1^\phi(t, \hat{X}(t), v_1, \hat{a}_2(t), \hat{b}_2(t), \hat{c}_2(t, \cdot)) \big| \mathcal{E}_t^{(1)} \right] \tag{61}$$
$$= \mathrm{E}\left[H_1^\phi(t, \hat{X}(t), u_1^*(t), \hat{a}_2(t), \hat{b}_2(t), \hat{c}_2(t, \cdot)) \big| \mathcal{E}_t^{(1)} \right]$$

  for all $t \in [0, T]$, a.s. Here $\hat{X}(t) = X^{u_1^*, \phi(u_1^*)}(t)$ and $\hat{a}_1(t), \hat{b}_1(t), \hat{c}_1(t, \cdot)$ is the solution of (55)-(56) corresponding to $u_1 = u_1^*$.

-

$$\mathrm{E}\left[\int_0^T (\hat{X}(t) - X^{u_1, \phi(u_1)}(t))^2 \left(\hat{b}_1^2(t) + \int_\mathbb{R} \hat{c}_1^2(t, \xi) \nu(d\xi) \right) dt \tag{62}$$
$$+ \int_0^T \hat{a}_1^2(t) \left(\sigma^2(t, X^{u_1, \phi(u_1)}(t), u_1(t), \phi(u_1(t))) \right.$$
$$+ \int_\mathbb{R} \gamma^2(t, X^{u_1, \phi(u_1)}(t), u_1(t), \phi(u_1(t)), \xi) \nu(d\xi) \right) dt \right] < \infty$$

  for all $u_1 \in \mathcal{A}_\mathcal{E}^{(1)}$.

Then $(u_1^*, \Phi(u_1^*)) \in \mathcal{A}_\mathcal{E}^{(1)} \times \mathcal{A}_\mathcal{E}^{(2)}$ is a Stackelberg equilibrium for the game.

Proof

Part i) is proved by applying the maximum principle for optimal stochastic control with respect to $u_2 \in \mathcal{A}_\mathcal{E}^{(2)}$ of the state process $X^{u_1, u_2}(t)$ for fixed $u_1 \in \mathcal{A}_\mathcal{E}^{(1)}$, as presented in Baghery and

26

Øksendal (2007). See also Framstad et. al (2004), Øksendal and Sulem (2007). Similarly, part ii) is proved by applying the same maximum principle for optimal stochastic control with respect to $u_1 \in \mathcal{A}_{\mathcal{E}}^{(1)}$ of the state process $X^{u_1, \phi(u_1)}(t)$, for the given function ϕ.

$\square$

8.1 Applications to the newsvendor problem

We now apply our general result for stochastic Stackelberg games to the newsvendor problem. In the newsvendor problem we have the control $u = (u_1, u_2)$ where $u_1 = w$ is the wholesale price, and $u_2 = (q, R)$ with q the order rate and R the retail price. Moreover $X_t = D_t$,

$$f_1(t, X(t), u(t)) = (w_t - M)q_t, \qquad\qquad g_1 = 0, \qquad (63)$$

$$f_2(t, X(t), u(t)) = (R_t - S)\min(D_t, q_t) - (w_t - S)q_t, \qquad \text{and} \qquad g_2 = 0. \qquad (64)$$

Therefore by (51)

$$H_2(t, D_t, q_t, R_t, w_t, a_2, b_2, c_2(\cdot)) = (R_t - S)\min(D_t, q_t) - (w_t - S)q_t \qquad (65)$$
$$+ a_2(t)\mu(t, D_t, R_t) + b_2(t)\sigma(t, D_t, R_t)$$
$$+ \int_{\mathbb{R}} \gamma(t, D_t, R_t, \xi)c_2(\xi)\nu(d\xi)$$

Similarly by (54) , with $u_2 = \phi(u_1) = (\phi_1(w), \phi_2(w)) = (q(w), R(w))$

$$H_1^{\phi}(t, D_t, w_t, a_1(t), b_1(t), c_1(t, \cdot)) \qquad (66)$$

$$= (w_t - M)\phi_1(w_t) + a_1(t)\mu(t, D_t, \phi_2(w_t)) + b_1(t)\sigma(t, D_t, \phi_2(w_t)) \qquad (67)$$

$$+ \int_{\mathbb{R}} c_1(t, \xi)\gamma(t, D_t, \phi_2(w_t), \xi)\nu(d\xi) \qquad (68)$$

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

$$dD_t = \mu(t, D_t, R_t)dt + \sigma(t, D_t, R_t)dB_t \tag{69}$$

$$+ \int_{\mathbb{R}} \gamma(t, D_t, R_t, \xi)\tilde{N}(dt, d\xi); \qquad t \in [0, T]$$

$$D_0 = d_0 \in \mathbb{R} \tag{70}$$

To find a Stackelberg equilibrium we use Theorem A.1. Hence by (57) we first maximize

$$\mathrm{E}\left[H_2(t, D_t, q_t, R_t, w_t, a_2(t), b_2(t), c_2(t, \cdot))\big|\mathcal{E}_t^{(2;j)}\right] \tag{71}$$

with respect to $u_2^{(j)}$, $j = 1, 2$, where $u_2^{(1)} = q, u_2^{(2)} = R$. The first order conditions give the equations

$$\mathrm{E}\left[(\hat{R}_t - S)\mathcal{X}_{[0,D_t]}(\hat{q}_t) - w_t + S\big|\mathcal{E}_t^{(2,1)}\right] = 0 \tag{72}$$

and

$$\mathrm{E}\left[\min(D_t, \hat{q}_t) + a_2(t)\frac{\partial \mu}{\partial R}(t, D_t, \hat{R})\right. \tag{73}$$

$$\left. + b_2(t)\frac{\partial \sigma}{\partial R}(t, D_t, \hat{R}) + \int_{\mathbb{R}} c_2(t, \xi)\frac{\partial \gamma}{\partial R}(t, D_t, \hat{R}, \xi)\nu(d\xi)\big|\mathcal{E}_t^{(2,2)}\right] = 0$$

for the optimal values $\hat{q}_t, \hat{R}_t$. Let $\hat{q}_t = \phi_1(w_t), \hat{R}_t = \phi_2(w_t)$ be the solution of this coupled system. Next, by (61) we maximize

$$\mathrm{E}\left[H_1^{\phi}(t, \hat{D}_t, w_t, a_1(t), b_1(t), c_1(t, \cdot))\big|\mathcal{E}_t^{(1)}\right] \tag{74}$$

$$= \mathrm{E}\left[(w_t - M)\phi_1(w_t) + a_1(t)\mu(t, D_t, \phi_2(w_t))\right.$$

$$\left. + b_1(t)\sigma(t, D_t, \phi_2(w_t)) + \int_{\mathbb{R}} c_1(t, \xi)\gamma(t, D_t, \phi_2(w_t), \xi)\nu(d\xi)\big|\mathcal{E}_t^{(1)}\right]$$

with respect to w_t. This gives the first order condition

$$(\hat{w}_t - M)\phi_1'(\hat{w}) + \phi_1(\hat{w}_t) + \phi_2'(\hat{w})\mathrm{E}\left[a_1(t)\frac{\partial\mu}{\partial R}(t, D_t, \phi_2(\hat{w}_t))\right. \tag{75}$$

$$\left. +b_1(t)\frac{\partial\sigma}{\partial R}(t, D_t, \phi_2(\hat{w}_t)) + \int_{\mathbb{R}} c_1(t, \xi)\frac{\partial\gamma}{\partial R}(t, D_t, \phi_2(\hat{w}_t), \xi)\nu(d\xi)\Big|\mathcal{E}_t^{(1)}\right] = 0$$

for the optimal value $\hat{w}_t$. The processes $a_i(t), b_i(t), c_i(t, \xi)$ are given by (52)-(53) for $i = 2$ and (55)-(56) for $i = 1$.

We summarize what we have proved in the following theorem.

Theorem A.2

Let $\hat{q}_t = \phi_1(w_t), \hat{R}_t = \phi_2(w_t)$ be the optimizers of (71). Assume that $\phi_i \in C^1$ and that the conditions of Theorem A.1 are satisfied. Let $\hat{w}_t$ be the optimizer of (74). Suppose

$$u^* = (\hat{w}_t, (\phi_1(\hat{w}_t), \phi_2(\hat{w}_t))) \in \mathcal{A}_{\mathcal{E}}^{(1)} \times \mathcal{A}_{\mathcal{E}}^{(2)}$$

Then u^* is a Stackelberg equilibrium for the newsvendor problem with state $X_t = D_t$ given by (69) and performance functionals

$$J_1(w, (q, R)) = \mathrm{E}\left[\int_\delta^T (w_t - M)q_t dt\right] \qquad \text{(manufacturer's profit)} \tag{76}$$

$$J_2(w, (q, R)) = \mathrm{E}\left[\int_\delta^T \left((R_t - S)\min(D_t, q_t) - (w_t - S)q_t\right)dt\right] \qquad \text{(retailer's profit)} \tag{77}$$

In other words

$$\max_{(q,R)\in\mathcal{A}_{\mathcal{E}}^{(2)}} \{J_2(w, (q, R))\} = J_2(w, (\phi_1(w), \phi_2(w))) \tag{78}$$

and

$$\max_{w\in\mathcal{A}_{\mathcal{E}}^{(1)}} \{J_1(w, (\phi_1(w), \phi_2(w)))\} = J_1(\hat{w}, (\phi_1(\hat{w}), \phi_2(\hat{w}))) \tag{79}$$

Remark

Note that if R is fixed and cannot be chosen by the retailer, then (73) is irrelevant and we are left with (72) leading to the much simpler equations that we reported in Theorem 3.2.1.

9 Appendix B

In this appendix we give complete proofs for all unproved statements given in Section 4 and 5. We start with a non-trivial estimate for the standard normal distribution, which will be crucial in the proofs of unique maxima.

Lemma B.1

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

$$h_m[z] = z(1 - m - G[z]) - G'[z] \tag{80}$$

Then

$$h_m[z] < 0 \qquad \text{for all } z \in \mathbb{R} \tag{81}$$

Proof

Note that if $z \geq 0$, then $h_m[z] \leq h_0[z]$ and if $z \leq 0$, then $h_m[z] \leq h_1[z]$. It hence suffices to prove the lemma for $m = 0$ and $m = 1$. Using $G''[z] = -x \cdot G'[z]$, it is easy to see that $h_m''[z] = -G'[z] \leq 0$. 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 $h_1[z]$ is strictly decreasing, and that $\lim_{z \to -\infty} h_1[z] = 0$. This proves that h_0 and h_1 are strictly negative, completing the proof of the lemma.

□

Details for Proposition 4.1.1

From (12), we easily see that the statement $q_t \leq D_t$ is equivalent to the inequality

$$q_t - \left(D_{t-\delta} e^{-a\delta} + \mu(1 - e^{-a\delta})\right) \leq \int_{t-\delta}^t \sigma e^{a(s-t)} dB_s \tag{82}$$

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

$$\mathrm{E}\left[\mathcal{X}_{[0,D_t]}(\hat{q}_t)|\mathcal{E}_t\right] = 1 - G\left[\frac{q_t - \left(D_{t-\delta}e^{-a\delta} + \mu(1 - e^{-a\delta})\right)}{\sqrt{\frac{\sigma^2(1-e^{-2a\delta})}{2a}}}\right] \tag{83}$$

and (13) follows trivially from (9). It remains to prove that the function Ψ_y has a unique maximum if $\Phi_y[M] > 0$. First put

$$\hat{y} = \frac{y \cdot e^{-a\delta} + \mu(1 - e^{-a\delta})}{\sigma\sqrt{\frac{1-e^{-2a\delta}}{2a}}} \tag{84}$$

and note that Ψ_y is proportional to

$$(w - M)\left(\hat{y} + G^{-1}\left[1 - \frac{w - S}{R - S}\right]\right) \tag{85}$$

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 get that Ψ_y is proportional to

$$(R - S)\left(1 - G[z] - \frac{M - S}{R - S}\right)(\hat{y} + z) \tag{86}$$

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

$$(1 - m - G[z])(\hat{y} + z) \tag{87}$$

$\Phi_y[M] > 0$ 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

$$f_m[z] = (1 - m - G[z])(\hat{y} + z) \qquad \text{on the interval } -\hat{y} \leq z \leq G^{-1}[1 - m] \tag{88}$$

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

$$f'_m[z] = -G'[z](\hat{y} + z) + (1 - m - G[z]) \tag{89}$$

Note that $f'_m[-\hat{y}] > 0$, and that $f_m[-\hat{y}] = f_m[G^{-1}[1-m]] = 0$. These functions hence 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

$$f''_m[z_0] = z_0(1 - m - G[z_0]) - 2G'[z_0] < z_0(1 - m - G[z_0]) - G'[z_0] < 0 \tag{90}$$

by Lemma B.1. The function is hence quasi-concave and has a unique, strictly positive maximum. That completes the proof of Proposition 4.1.1.

$\square$

Details for Proposition 4.2.1

In this case the statement $q_t \le D_t$ is equivalent to the inequality

$$q_t - \left(D_{t-\delta} e^{-\int_{t-\delta}^{t} a(u)du} + \int_{t-\delta}^{t} a(s)\mu(s)e^{-\int_{t-\delta}^{t} a(u)du} \right) \le \int_{t-\delta}^{t} \sigma(s)e^{-\int_{t-\delta}^{t} a(u)du}dB_s \tag{91}$$

Again the left hand side is $\mathcal{E}_t$-measurable, while the right hand side is normally distributed and independent of $\mathcal{E}_t$. The calculations in the proof of Proposition 4.1.1 can be repeated line by line for each fixed value of t, proving the general case in Proposition 4.2.1.

$\square$

Details for Proposition 5.1.1

From (22), we easily see that the statement $q_t \le D_t$ is equivalent to the inequality

$$\ln\left[\frac{q_t}{D_{t-\delta}}\right] - (a - \frac{1}{2}\sigma^2)\delta \le \sigma(B_t - B_{t-\delta}) \tag{92}$$

The left hand side is $\mathcal{E}_t$-measurable, while the right hand side is normally distributed and independent of $\mathcal{E}_t$. It is then straightforward to prove that

$$\mathrm{E}\left[\mathcal{X}_{[0,D_t]}(\hat{q}_t)|\mathcal{E}_t\right] = 1 - G\left[\frac{\ln\left[\frac{q_t}{D_{t-\delta}}\right] - (a - \frac{1}{2}\sigma^2)\delta}{\sqrt{\sigma^2\delta}}\right] \tag{93}$$

Hence it follows from (9) that

$$q_t = D_{t-\delta} \cdot \exp\left[(a - \frac{1}{2}\sigma^2)\delta + \sqrt{\delta\sigma^2} \cdot G^{-1}\left[1 - \frac{w - S}{R - S}\right]\right] \tag{94}$$

With this order quantity, the expected profit for the manufacturer is

$$\mathrm{E}[D_{t-\delta} \cdot (w_t - M)\exp\left[(a - \frac{1}{2}\sigma^2)\delta + \sqrt{\delta\sigma^2} \cdot G^{-1}\left[1 - \frac{w_t - S}{R - S}\right]\right] \tag{95}$$

In general w_t can be a random variable. If $w^* = \mathrm{Argmax}[\Psi]$, where

$$\Psi[w] = (w - M)\exp\left[(a - \frac{1}{2}\sigma^2)\delta + \sqrt{\delta\sigma^2} \cdot G^{-1}\left[1 - \frac{w - S}{R - S}\right]\right] \tag{96}$$

we have, however, that

$$\mathrm{E}\left[D_{t-\delta} \cdot (w_t - M)\exp\left[(a - \frac{1}{2}\sigma^2)\delta + \sqrt{\delta\sigma^2} \cdot G^{-1}\left[1 - \frac{w_t - S}{R - S}\right]\right]\right] \le \mathrm{E}[D_{t-\delta}] \cdot \Psi[w^*] \tag{97}$$

with equality if $w_t = w^*$. Hence w^* is optimal. It remains to prove that $\mathrm{Argmax}[\Psi]$ is unique. If we put $b = \sqrt{\delta\sigma^2}$, it follows that Ψ is proportional to a function of the form

$$w \mapsto (w - M)\exp\left[b\,G^{-1}\left[1 - \frac{w - S}{R - S}\right]\right] \tag{98}$$

where $b > 0$. Make a monotone change of variables using $z = G^{-1}\left[1 - \frac{w-S}{R-S}\right]$. With this change of variables we get that Ψ is proportional to

$$(R - S)\left(1 - G[z] - \frac{M - S}{R - S}\right)\exp[b\,z] \tag{99}$$

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

$$(1 - m - G[z])\exp[b\,z] \tag{100}$$

For each fixed $0 \le m \le 1, b > 0$ consider the function

$$f_m[z] = (1 - m - G[z])\exp[b\,z] \qquad \text{on the interval } -\infty < z \le G^{-1}[1 - m] \tag{101}$$

33

We have

$$f'_m[z] = -G'[z]\exp[b\,z] + (1 - m - G[z])b\exp[b\,z] \tag{102}$$

Note that $\lim_{z\to-\infty} f_m[z] = 0$, $f_m[G^{-1}[1-m]] = 0$, and $f'_m[G^{-1}[1-m]] < 0$. The function hence has at least one strictly positive maximum. To see that the maximum is unique, find z_0 s.t. $f'_m[z_0] = 0$. Using $G''[z] = -z \cdot G'[z]$, we can simplify the expression to get

$$f''_m[z_0] = (z_0 - b)G'[z_0]\exp[b\,z] \tag{103}$$

From Lemma B.1 and $f'_m[z_0] = 0$ we get

$$(1 - m - G[z_0])b = G'[z_0] > (1 - m - G[z_0])z_0 \tag{104}$$

Hence if $f'_m[z_0] = 0$, we must have $z_0 < b$, which implies $f''_m[z_0] < 0$. The function is hence quasi-concave and has a unique, strictly positive maximum. That completes the proof of Proposition 5.1.1.

$\square$

Details for Proposition 5.2.1

In the case with variable coefficients, we have

$$D_t = D_{t-\delta} \cdot \exp\left[\int_{t-\delta}^{t} \mu(s) - \frac{1}{2}\sigma^2(s)ds + \int_{t-\delta}^{t}\sigma(s)dB_s\right] \tag{105}$$

Put

$$\hat{\mu}(t) = \int_{t-\delta}^{t} \mu(s) - \frac{1}{2}\sigma^2(s)ds \qquad \hat{\sigma}^2(s) = \int_{t-\delta}^{t}\sigma^2(s)ds \tag{106}$$

Since the exponent in (105) is normally distributed and independent of $\mathcal{E}_t$, we get

$$\mathrm{E}\left[\mathcal{X}_{[0,D_t]}(\hat{q}_t)|\mathcal{E}_t\right] = 1 - G\left[\frac{\ln\left[\frac{q_t}{D_{t-\delta}}\right] - \hat{\mu}(t)}{\hat{\sigma}(t)}\right] \tag{107}$$

Hence it follows from (9) that

$$q_t = D_{t-\delta} \cdot \exp\left[\hat{\mu}(t) + \hat{\sigma}(t) \cdot G^{-1}\left[1 - \frac{w - S}{R - S}\right]\right] \tag{108}$$

34

With this order quantity, the expected profit for the manufacturer is

$$
\mathrm{E}\left[D_{t-\delta} \cdot (w_t - M) \exp\left[\hat{\mu}(t) + \hat{\sigma}(t) \cdot G^{-1}\left[1 - \frac{w-S}{R-S}\right]\right]\right] \tag{109}
$$

The calculations in the proof of Proposition 5.1.1 can now be repeated line by line for each fixed t proving the general case in Proposition 5.2.1. $\qquad\square$

Details for subsection 5.3

From (32) it follows that

$$
D_t = D_{t-\delta} \exp\left[\int_{t-\delta}^{t}\left(B_s(\omega) - \frac{1}{2}\sigma^2(s) + \int_{\mathbb{R}_0}\log[1 + \gamma(s,z)] - \gamma(s,z)\nu(dz)\right)ds \right. \tag{110}
$$
$$
\left. + \int_{t-\delta}^{t}\sigma(s)dB_s + \int_{t-\delta}^{t}\int_{\mathbb{R}_0}\log[1 + \gamma(s,z)]\tilde{N}(ds,dz)\right]
$$
$$
= D_{t-\delta} \cdot \exp\left[\int_{t-\delta}^{t}B_s(\omega)ds\right] \tag{111}
$$
$$
\cdot \exp\left[\int_{t-\delta}^{t}\left(-\frac{1}{2}\sigma^2(s) + \int_{\mathbb{R}_0}\log[1 + \gamma(s,z)] - \gamma(s,z)\nu(dz)\right)ds \right. \tag{112}
$$
$$
\left. + \int_{t-\delta}^{t}\sigma(s)dB_s + \int_{t-\delta}^{t}\int_{\mathbb{R}_0}\log[1 + \gamma(s,z)]\tilde{N}(ds,dz)\right]
$$

The problem here is the second term $\exp\left[\int_{t-\delta}^{t}B_s(\omega)ds\right]$ which is *not* independent of $\mathcal{E}_t$. Integrate by parts to see that

$$
\exp\left[\int_{t-\delta}^{t}B_s(\omega)ds\right] = \exp\left[\delta B_{t-\delta} + t(B_t - B_{t-\delta}) - \int_{t-\delta}^{t}sdB_s\right] \tag{113}
$$

from which it follows that

$$
D_t = D_{t-\delta} \cdot \exp\left[\delta B_{t-\delta}\right] \cdot X_t \tag{114}
$$

where X_t is given by (34). Here the first two terms are $\mathcal{E}_t$-measurable, while the last term is independent of $\mathcal{E}_t$. It is then straightforward to see that (35) follows from (9).

$\qquad\square$

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 SP. Sethi 2007. A multiperiod newsvendor problem with partially observed demand. *Mathematics of Operations Research*, Vol. 23, No.2, 322–344.

[3] Bensoussan A., M. Çakanyıldırım and SP. Sethi 2009. Optimal Ordering Policies for Stochastic Inventory Problems with Observed Information Delays, *Production and Operations Management*, Vol. 18, Issue 5, 546-559.

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

[5] Cachón GP., 2003. Supply Chain Coordination with Contracts. Chapter 6 in The Handbook of Operations Research and Management Science: Supply Chain Management: Design, Coordination and Operation. A. G. de Kok and S. C. Graves (eds), Elsevier.

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

[7] 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.

[8] Gallego G. and I. Moon, 1993. The distribution free newsboy problem: Review and extensions. *Journal of the Operational Research Society* 44, No.8, 825–834.

[9] Kogan K. and S. Lou, 2003. Multi-stage newsboy problem: A dynamical model. *European Journal of Operational Research* 149, No. 2, 448–458.

[10] Kogan K., 2003. Scheduling parallel machines by the dynamic newboy problem. *Computers & Operations Research* 31, No. 3, 429–443.

[11] Matsuyama K., 2004. The multi-period newsboy problem. *European Journal of Operational Research* 171, No. 1, 170–188.

[12] Øksendal B. and A. Sulem, 2007. *Applied Stochastic Control of Jump Diffusions.* Second Edition. Springer.

[13] Perakis G. and G. Roels, 2008. Regret in the newsvendor model with partial information. *Operations Research* 56, No. 1, 188-203.

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

[15] 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), pp. 201–209.

[16] 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.