

Geometric insights and proofs on optimal inventory control policies

Nicky D. van Foreest, Onur. A. Kilic

Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands
{n.d.van.foreest@rug.nl, o.a.kilic@rug.nl}

We develop a unifying framework to prove the existence of optimal policies for a large class of inventory systems. The framework is based on the transformation of the inventory control problem into a game, each round of which corresponds to a single replenishment cycle. By using parametrized optimization methods we show that finding the equilibrium in this game is equivalent to finding the optimal long-run average cost of the inventory problem, and demonstrate that any root-finding algorithm can be used to find the optimal inventory policy. We then cast the associated parametrized problem into an optimal stopping problem. The analysis of this problem reveals that the structure of the optimal policy is a geometrically obvious consequence of the shape of the inventory cost function. It also allows us to derive bounds on the optimal policy parameters and devise a straightforward scheme to efficiently compute them. The proposed framework enjoys the power tools of optimal stopping theory and thereby provides a new and powerful line of attack that can address periodic and continuous review inventory problems with discrete and continuous demands in a unified fashion.

Key words: optimal stochastic control, inventory theory, dynamic programming, free boundary problems

History:

1. Introduction

As one of the most fruitful application areas of operations research, inventory control has attracted an enormous amount of research interest in the past decades. A large variety of inventory control problems have been analyzed which are characterized by different assumptions on demand (e.g., continuous or discrete, stochastic or deterministic), service discipline (e.g., backordering or lost sales), and cost structure (e.g., linear and/or fixed replenishment costs, convex or quasi-convex inventory costs). Also, a variety of inventory control policies, that is, policies that determine when

and how much to order, have been introduced.

The literature on inventory control has three focal points: proving the existence of an optimal policy under a certain criterion and in some appropriate policy space, establishing insight into the structure of an optimal policy, and, finally, computing such an optimal policy. Despite that the work on inventory control has resulted in many powerful insights on these three points, the models and methods are typically ad-hoc, that is, each problem tends to require its own methods for the existence proofs and specific algorithms for the computational aspects. This is an unfortunate state of affairs, as most inventory control problems start from similar inventory dynamics. A comprehensive, all-encompassing approach to analyze the theoretical and numerical aspects of all these inventory systems simultaneously is therefore in dire need. This paper aims at serving this purpose.

1.1. Overview and contribution

In this paper, we develop a unifying framework by which we can prove the existence of optimal policies and compute efficiently such policies for a large class of inventory systems, including all single-item and single-location inventory control problems with constant lead times in which the inventory process can be modeled as a Markov process. For the rest, the problem can be very general: the inventory system can be under periodic or continuous review, the demand can be discrete or continuous, demand not satisfied from stock can be backlogged or lost. There is a fixed cost K per replenishment. The expected holding and backlogging costs accrued in response to the realized demand (during the lead time) are captured by a non-negative expected cost function $L(\cdot)$. We establish very general conditions on $L(\cdot)$ under which optimal policies exist. For the proofs that (s, S) -policies are optimal we assume that $L(\cdot)$ is quasi-convex, for the numerical procedures this assumption is unnecessary. We focus on optimality under the long-run average cost criterion. Because procurement costs are immaterial to the average cost, we do not take these costs into account.

We next sketch our approach. The core idea is to avoid at first the infinite-horizon inventory control problem, and to consider a single replenishment cycle instead. A replenishment cycle starts

at time 0 with an initial inventory level $I_0 = S$. Then, demand is served until the process is stopped at some time τ . Over the replenishment cycle, costs accrue following $L(I_t)$, $t \in [0, \tau)$, and the ordering cost K is incurred as terminal cost. Next, we introduce a reward $g > 0$ per unit time which is offered as compensation for the ordering and inventory holding cost. We are interested in the problem of finding an initial inventory level S and a stopping time τ such that the expected cycle cost $W^g(S, \tau) = K + \mathbb{E}_S \left[\int_0^\tau (L(I_t) - g) dt \right]$ is minimized. We make at this point the crucial observation that the optimization over τ is an optimal stopping problem, and that the optimal stopping time τ^g can be characterized as the hitting time of a set that only depends on the inventory cost function L and the reward g . The proof of the existence of an optimal stopping time τ^g will then imply the existence of a long-run average optimal inventory control policy. This interpretation of the inventory control problem turns out to be particularly fruitful; we discuss its advantages below.

The first advantage of the proposed approach is that it leads to an intuitive graphical interpretation of the problem. It is clear (intuitively at least) from the graph of $L(\cdot) - g$ that the cycle cost $W^g(S, \tau)$ can be minimized by restricting the inventory process to the set where $L(\cdot) - g$ is negative. Hence, if L is quasi-convex, we certainly stop the process at the first moment that the inventory position becomes smaller than the left root of $L(\cdot) - g$. It is also quite evident that it cannot be optimal to initiate a replenishment cycle at an inventory position larger than the right root of $L(\cdot) - g$, because there we have $L(\cdot) - g > 0$ and including such terms in the cost functional W^g will only add to the cycle cost. As a consequence, the graph of $L(\cdot) - g$ suggests that the optimal stopping time can be characterized as an optimal stopping set, i.e., the set of points smaller than the left root of $L(\cdot) - g$, and the optimal starting level is smaller than or equal to the right root. In terms of the inventory control problem, the optimal stopping set is the optimal reorder set, i.e., once the inventory hits this set, it is optimal to place a replenishment. The fact that the left and right roots of $L(\cdot) - g$ confine the search space for the starting level S and stopping time τ is not just numerically relevant, it also has crucial ramifications for the structure of the optimal inventory policy. If the optimal stopping set is a connected set of the type $(-\infty, s]$ for every g , then

it must follow that the optimal inventory policy is of (s, S) -type. As we will see, the optimality of (s, S) -policies then becomes a geometrically obvious consequence of the shape of the cost function $L(\cdot)$.

The second advantage of our approach is that the existence of an optimal inventory policy follows directly from the existence of the optimal stopping times $\{\tau^g\}$. This opens up the possibility to study inventory control problems within a broader set of control problems where the “power tools” of optimal stopping theory become available for existence proofs. By means of the optimal stopping theory, we provide three new proofs of the optimality of (s, S) -policies. The first of these applies to periodic-review systems with discrete demand and does not hinge on specific functional properties such as K -convexity, but relies solely on dynamic programming concepts which were available when Scarf (1959) and Iglehart (1963) provided their first optimality proofs. Due to its intuitive structure and easy mathematical background, we believe that it has the potential to become a standard textbook proof of the optimality (s, S) -policies. The second proof also applies to periodic-review systems with discrete demand. It uses linear programming and duality theory, and has considerable appeal by its geometric nature. The third proof applies to continuous-review systems with discrete or continuous demand, and is based on formulating the optimal stopping problem as a free-boundary problem, c.f., Peskir and Shiryaev (2006). The related theory provides a new and powerful line of attack to address such continuous-review inventory systems with a continuous state-space. We note that—perhaps contrary to common belief—it is by no means a easy task to prove that continuous systems can be obtained as simple limits of their discrete counterparts, c.f., Shiryaev (2007). We refer to Beyer et al. (2010, Chapter 9) for a very interesting discussion and historical overview of the efforts that went into existence proofs for the continuous time- and state-space systems, and the traps these approaches had to overcome (or fell into). Of course, technical issues also remain with our approach, but what needs to be proven follows a clear and generic program, which can be applied to the analysis of many more inventory problems.

The third advantage of the proposed approach is its numerical appeal. The methods based on dynamic programming, duality theory, and free-boundary theory formulate the associated optimal

stopping problem in terms of concepts that follow straightaway from the dynamics of the controlled inventory process and the costs $L(\cdot)$ and K .

In summary, our approach seamlessly unite the three main aspects of inventory control: existence proofs, derivation of structural properties of optimal policies, and fast numerical methods to actually compute these optimal policies.

1.2. Related literature

The work on stochastic inventory control problems has had a significant impact on supply chain management practices well over half a century. Because it is far too ambitious to provide a comprehensive review of such a long-standing body of literature, here we opt for mentioning some key results. Extensive discussions of the classical and, probably, most relevant single-item inventory models and the control problems thereof can be found in, for example, Hadley and Whitin (1963), Zipkin (2000), and Axsäter (2006).

The current paper relates to the work on characterizing the structure of optimal control policies as well as computational procedures. Therefore, it is closely related to these two streams of research.

The challenge of establishing the structure of optimal inventory policies has been a center of interest in the inventory control literature. The (s, S) -policy in this context is of particular importance as it has been shown to be optimal under rather general conditions. Scarf (1959), Zabel (1962), Iglehart (1963) and Veinott and Wagner (1965) are the pioneering studies on the optimality of the (s, S) -policy for periodic-review systems. Beckmann (1961), Bather (1966), and Hordijk and Van der Duyn Schouten (1986), on the other hand, present seminal results for continuous-review systems. We refer the reader to Presman and Sethi (2006) and Beyer et al. (2010) for a discussion on the development of the optimality proofs of (s, S) -policies. The contemporary work on the characterization of optimal control policies such as Zheng (1991), Chen and Feng (2006), and Huh et al. (2011) provide alternative proofs of the optimality of the (s, S) -policy. The current paper falls into this line of research by providing new optimality proofs. These follow a different approach as compared to the earlier contributions, as they use optimal stopping theory to address the inventory control problem and make use of the structural properties thereof.

The literature on the computational aspects of optimal inventory control policies is rather limited. Despite the early developments at the theoretical frontier around 1960, it was not until the early 1990s that straightforward and reliable methods were introduced for computing the optimal parameters of well-known inventory control policies, rather than using straight enumeration. The contributions in this domain of research mostly concentrate on periodic review systems. We can mention, among others, Zheng and Federgruen (1991), Feng and Xiao (2000), and Chen and Feng (2006) where efficient algorithms are presented for computing optimal (s, S) -policies. We contribute this line of literature by providing a generic computational approach that applies to both discrete- and continuous-review systems. The methods we use to solve embedded optimal stopping problems follow straightaway from the dynamics of the controlled inventory process which in fact boil down to integrating the cost function over a finite interval, and the rest of the procedure involves nothing but a root finding problem.

The core of our approach is to cast the inventory control problem into a parametric optimization problem, following the highly useful idea of embedding optimization problems into higher-dimensional spaces (see e.g. Rockafellar 1974). We then show that the parametric problem can be modelled and solved as an optimal stopping problem. Van Foreest and Wijngaard (2014) has recently used a similar approach to prove the optimality of (s, S) -policies in continuous-review production-inventory systems. Their approach, however, does not extend to inventory systems, or, to be more precise, systems with jumps in the replenishment process. To cope with this, one needs to take the limit of the production rate to infinity, and then show that the entire analysis carries over to this limiting case. The easiest way to do this seems to be to use the standard theory of Markov processes (mainly the concept of ‘generator’), and interpret the ensuing optimal stopping problem as a free-boundary problem. This is in fact what we do in the current study.

1.3. Outline

The remainder the paper is organized as follows. In Section 2, we formally define the inventory system and the associated inventory control problem. In Section 3, we cast the inventory control

problem into a parametric optimization problem and discuss a simple and intuitive interpretation to understand of their equivalence. In Section 4, we provide an analysis of the parametric optimization problem and establish its relationship with an embedded optimal stopping problem. In Section 5, we address the optimal stopping problem associated with a periodic review inventory system with discrete demand. The analysis presented here results in two new proofs of the optimality of (s, S) -policies, the first based on dynamic programming, and the second on linear programming. In Section 6, we consider a continuous-review inventory system with compound Poisson demand. Once again, we approach the inventory control problem in terms of a parametric optimal stopping framework, yet we deal with this more general case by using free-boundary theory to formulate and solve the associated optimal stopping problems. This also leads to a new proof of the optimality of (s, S) -policies for continuous-review inventory control systems. In Section 7 we discuss a number of extensions to which our work directly generalizes.

2. Inventory model and problem definition

Consider an inventory system in which demands arrive according a counting process (N_t) such that N_t is the number of arrivals during the interval $[0, t]$. Let (T_n) denote the sequence of arrival epochs, i.e., $T_n = \inf\{t : N_t \geq n\}$. At time T_n an order arrives of size Y_n . The sequence of demand sizes (Y_n) forms a set of independent random variables all distributed as the common non-negative random variable Y with distribution function $F(x) = \mathbb{P}(Y \leq x)$. Hence, the total demand up to time t becomes $Z_t = \sum_{k=1}^{N_t} Y_k$.

When the reviews of the inventory occur periodically, we take $N_t = \lfloor t \rfloor$, i.e., the largest integer smaller than t , and we allow the demand Y_k occurring in k th periodic to be zero, i.e., $\mathbb{P}(Y_k = 0) > 0$. In case of continuous review, we take N_t to be Poisson distributed with mean λt , so that (Z_t) becomes a compound Poisson process. Clearly, in this case it makes no sense (but also does not harm) to include the occurrence of zero-sized demands.

An admissible inventory control policy π is given by a sequence of pairs of random variables $(\sigma_k, y_k), k = 1, 2, \dots$, where y_k is the amount ordered at replenishment time σ_k . It is natural to

assume that order sizes are positive, i.e., $y_k > 0$, and the replenishment times σ_k increase, i.e., $\sigma_{k-1} \leq \sigma_k$ for all k . The duration of the k th replenishment cycle is given by $\tau_k = \sigma_k - \sigma_{k-1}$. To prevent ‘look-aheads’ into the future, we require π to be non-anticipative, that is, the times (σ_k) are stopping times with respect to the filtration (\mathcal{F}_t) generated by (Z_t) , and each $y_k \in \mathcal{F}_{\sigma_k}$.

Supposing that the inventory starts at $I_0 = x$, the (right-continuous with left limits) sample paths of the inventory process I^π under the control π are given by the map

$$t \rightarrow I_t^\pi = x + \sum_{n=1}^{\infty} y_n \mathbf{1}\{\sigma_n \leq t\} - Z_t,$$

with $\mathbf{1}\{\cdot\}$ the indicator function.

It is easy to see that an (s, S) policy is admissible by taking the replenishment times $(\sigma_k) \subset (T_n)$ and (y_k) as follows. Whenever the inventory just prior to ordering $I_{T_{n-1}} - Y_n \leq s$, we place an order of size $y_k = S - (I_{T_{n-1}} - Y_n)$ at time $\sigma_k = T_n$. Otherwise, do not order. In a similar way we can construct the sample paths of the inventory process under any stationary policy.

Note that if the replenishment lead time is positive and constant, we use the inventory position in π rather than the inventory level.

Let $h(\cdot)$ be the inventory (holding and shortage) cost function. In case the replenishment lead time is positive, the expected cost is given by

$$L(x) = \mathbb{E} [h(x - \bar{Y})]$$

where \bar{Y} is the demand during the lead time. We henceforth write $L(\cdot)$ to cover the inventory cost whether the lead time is zero or not.

We say that a function h is *positive and (eventually) sufficiently large* if $h(0) = 0$, $h(\cdot) \geq 0$ and $h(x)$ becomes so large when $|x| \rightarrow \infty$ that it is not optimal to let the inventory drift to $\pm\infty$, either by not ordering at all or by ordering too much on average. A simple, but certainly not necessary, condition for this is $h(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. More generally, consider the function $\tilde{L}(x) = \mathbb{E}[L(x - Y)]\mathbb{E}[T] - K$. If $\tilde{L}(x) > 0$ for all $x < m$ with $m \ll 0$, it is eventually cheaper to place an

order than wait for the next arrival, and if $\tilde{L}(x) > 0$ for all $x > M$ with $M \gg 0$, it is not optimal to continuously place replenishments such that the inventory drifts to $+\infty$.

For the structure proofs we shall assume that h is such that L is quasi-convex and sufficiently large; if h is convex and sufficiently large this is guaranteed. Below, when we write that L is quasi-convex, we will always implicitly assume L is also eventually sufficiently large. Note that for the computational procedure to work, L only needs to be sufficiently large.

Finally, each replenishment costs a fixed amount $K > 0$, independent of the replenishment size.

We write the total cost up to time t under policy π as

$$G_t(x, \pi) = \mathbb{E}_x \left[\int_0^t L(I_s^\pi) ds + K \sum_{k=1}^{\infty} \mathbf{1}\{\sigma_k \leq t\} \right],$$

where $\mathbb{E}_x[\cdot]$ is the expectation of functionals of I^π such that $I_0 = x$. Define the long-run average expected cost under policy π as

$$G(x, \pi) = \limsup_{t \rightarrow \infty} \frac{G_t(x, \pi)}{t}.$$

In this paper, we are concerned with the optimization problem

$$G^* = \inf_{\pi} G(x, \pi) \tag{1}$$

where π runs over the space of admissible policies. Our aim is to show that there exists a *stationary* policy π^* in the (larger) class of admissible policies such the infimum is attained and that G^* is independent of the initial inventory level x , i.e., there exists π^* such that $G^* = G(x, \pi^*)$ for all x . Moreover, once we have proven that an optimal policy π^* exists, we are interested in identifying its structural properties, e.g., (s, S) , and establishing a fast scheme to compute it.

3. Parametric optimization with a reward

In this section we discuss an alternative, parametric, optimization problem by which we can find an optimal policy for the inventory control problem described in the previous section. This parametric problem, to be derived in Section 3.1, consists of an iteration of games, each game involving an optimal stopping problem. It turns out that the sequence of games converges to a ‘limit game’,

which provides an optimal inventory policy under the long-run average cost criterion. In Section 3.2 we relate this limit game to the inventory control problem (1). In Section 3.3 we discuss some computational issues related to actually finding the optimal policy.

3.1. Parametric cycle cost

We first introduce the concept of g -revised cycle cost. Perhaps the easiest way to explain this concept is by means of a game between a franchiser and a franchisee. The franchiser is responsible of serving demands (Y_n) over time and contemplates to relegate this service to a franchisee. The franchisee can start the game with any amount of inventory she chooses. Then, for the time she stays in the game, she receives a reward $g > 0$ per unit time, and pays $L(I_t)$ per unit time if the inventory level is I_t at time t . The franchisee can stop the game any time she likes, but whenever she chooses to stop, she pays a terminal cost K to replenish the inventory. We shall address the following questions: Considering the fixed participation fee K , the inventory cost function $L(\cdot)$, and the reward g per unit time, should the franchisee participate in this game, and if so what is the optimal policy?

Let us assume that the franchisee indeed accepts the offer and starts the game by setting her inventory level to $I_0 = x$. The franchisee's problem then reduces to deciding when to stop the process. Suppose she would use stopping rule τ , then the expected cost of starting at level x becomes

$$W^g(x, \tau) = K + \mathbb{E}_x \left[\int_0^\tau (L(I_s) - g) \, ds \right]. \quad (2)$$

Obviously, for given x , the franchisee will try to choose a stopping time τ such that her loss until the end of the game is minimized, in other words, she will try to solve

$$W^g(x) = \inf_{\tau} W^g(x, \tau) = K + \inf_{\tau} \mathbb{E}_x \left[\int_0^\tau (L(I_s) - g) \, ds \right]. \quad (3)$$

Here we make the crucially important observation that this is an *optimal stopping problem*. We assume for now that a minimizing policy, i.e., an optimal stopping time, exists for this problem, and postpone the associated technical issues to Section 5 and Section 6.

Once the optimal stopping time τ^g is known, the franchisee of course will also want to optimize over the inventory starting level x , that is, the franchisee aims to solve

$$W^g = \inf_x W^g(x). \quad (4)$$

The franchisee should therefore address the combined problem of finding an optimal stopping time τ^g and an optimal starting inventory level S^g solving (3) and (4), respectively; such that $W^g = W^g(S^g, \tau^g)$. We shall refer to this problem as the *g-revised problem*.

We now consider the franchisee's perspective on the value of W^g . If $W^g > 0$, the franchisee cannot obtain a positive expected profit under any policy, not even under the optimal policy (S^g, τ^g) . Thus, provided rationality, she will decide not to participate in the game. If, however, $W^g < 0$, the policy (S^g, τ^g) will provide a positive expected reward (equal to $-W^g$). Thus, she will decide to accept the offer. From the franchiser's perspective; if the franchisee is eager to participate in the game, then the compensation g must be (slightly) higher than needed, and in a next negotiating round the franchiser can lower the reward. Therefore, in the course of a negotiation process between franchiser and franchisee, the reward will eventually settle to a value g^* such that $W^{g^*} = 0$.

The following lemma introduces some important properties of W^g which we make heavy use of in our analysis.

LEMMA 1. *If there exists a policy (S^g, τ^g) that solves (3) and (4) for any $g \geq 0$, then the following hold.*

- (a) *W^g is decreasing in g .*
- (b) *There exist g^- and g^+ such that $W^{g^-} \leq 0 \leq W^{g^+}$.*
- (c) *W^g is continuous in g .*

With this lemma we can establish that an equilibrium exists in the game between the franchiser and franchisee. By applying the intermediate value theorem to the function $g \rightarrow W^g$ the following theorem immediately follows.

THEOREM 1. *There exists a g^* such that $W^{g^*} = 0$.*

The solution g^* of $W^{g^*} = 0$ has an interesting interpretation. Let us write $(S^*, \tau^*) = (S^{g^*}, \tau^{g^*})$ for the policy associated to g^* . Then, by (3) and (4), we have

$$0 = W^{g^*} = K + \mathbb{E}_{S^*} \left[\int_0^{\tau^*} L(I_s) ds \right] - g^* \mathbb{E}_{S^*} [\tau^*]$$

which implies

$$g^* = \frac{K + \mathbb{E}_{S^*} \left[\int_0^{\tau^*} L(I_s) ds \right]}{\mathbb{E}_{S^*} [\tau^*]}. \quad (5)$$

The meaning of g^* is evident from this expression: it is precisely the franchisee's average expected replenishment cycle cost under the policy (S^*, τ^*) . Moreover, since two consecutive games are independent of each other, the franchisee can use the policy (S^*, τ^*) cycle after cycle. Therefore, by the renewal reward theorem, g^* is also the long-run average expected cost per unit time under the policy (S^*, τ^*) .

In fact, Lemma 1 and Theorem 1 have further spin off. In the next section we use these to prove that an optimal stationary policy exists, and in Section 3.3 we discuss how Lemma 1 can be used in the numerical computation of optimal policies.

3.2. Relation to the inventory control problem

The interpretation of Lemma 1, in particular part (b), and Theorem 1 provide the key insight to understand that a long-run average cost minimizing policy exists. By combining this with the structure of the stopping times τ^g , we can show that these optimal policies are stationary, i.e., only dependent only on the inventory level.

THEOREM 2. *The solution (S^*, τ^*) of the g -revised problem (4) is an optimal stationary inventory policy π^* and its average cost $g^* = G(x, (S^*, \tau^*))$ for all x in (1).*

Proof. Let us start with stationarity. For this we use the result of Section 5 and Section 6 that for given g , the optimal stopping time in (3) can be characterized as a hitting time

$$\tau^g = \inf\{t \geq 0 : I_t \in \mathcal{D}^g\},$$

and the hitting set \mathcal{D}^g does *not* depend on the inventory level at which the cycle starts. This implies that the stopping time in $W^g(x)$ in (3) does not depend on x . Moreover, no matter where this

stopping set is entered, it follows from (4) that we prefer to start at level S^g . Thus, the policy that minimizes the g -revised cycle cost is stationary. In particular, this is true for g^* . Hence, (S^*, τ^*) is also a stationary policy.

For the optimality, we consider two cases. First, suppose that g is such that $W^g > 0$. Then, by following the reasoning that lead to (5), we obtain that

$$g < \frac{K + \mathbb{E}_{S^g} \left[\int_0^{\tau^g} L(I_s) ds \right]}{\mathbb{E}_{S^g} [\tau^g]}.$$

In words, g is smaller than right hand side which, by the renewal reward theorem, is the long run average cost of the policy (S^g, τ^g) . From (3) and (4) we know that (S^g, τ^g) minimizes the expected cycle g -revised cost. It must then follow that there is no policy, not even (S^g, τ^g) , that can yield g as a long-run average reward.

Next, suppose that g is such that $W^g < 0$. Then, we obtain

$$g > \frac{K + \mathbb{E}_{S^g} \left[\int_0^{\tau^g} L(I_s) ds \right]}{\mathbb{E}_{S^g} [\tau^g]}.$$

This means, once again, by the renewal reward theorem, that the long-run average cost of the policy (S^g, τ^g) is smaller than g . Hence, there exists a policy with lower cost than g .

By combining the two observations we conclude that $(S^*, \tau^*) = (S^{g^*}, \tau^{g^*})$ is an optimal stationary policy, and we have found a solution of the inventory control problem (1).

Another way to obtain the same insight is as follows. Clearly, (S^*, τ^*) minimizes the expected cost of one cycle. But then, by using this policy cycle after cycle, we minimize the cost cycle after cycle, hence (S^*, τ^*) must be a long-run average cost minimizing policy. \square

It is of paramount importance to realize that the original inventory problem (1) and the g -revised problem (4) are related but not the same. The main difference lies in what is being optimized. In the original problem, the aim is to find a policy π^* that minimizes the long-run expected average cost per cycle. The standard way of approaching this problem is to find a sequence of policies (π_k) with decreasing long-run expected average costs per cycle such that $\pi_k \rightarrow \pi^*$. This is challenging as it requires to prove the existence of a limiting policy π^* . Also, the structure of this limiting policy

needs to be addressed separately as this is not necessarily the same as the structure of the policies (π_k) . We refer to (Beyer et al. 2010, Chapter 9) for an in depth discussion of the subtle issues that need to be accounted for with this approach. In the g -revised problem, on the other hand, the objective is to find a policy (S^g, τ^g) that minimizes the expected total loss $W^g = W^g(S^g, \tau^g)$ over a replenishment cycle, given a reward rate g . The optimal policy minimizing the long-run average cost can then be obtained in the context of the parametric optimization problem by finding a g^* such that $W^{g^*} = 0$. This approach has the pivotal advantage that the existence and the structure of the limiting policy immediately follow those of the corresponding g -revised problems—which happens to be very easy to identify as we illustrate in the following sections.

3.3. Computational aspects

Lemma 1 and Theorem 1 provide some useful properties in the computation of the limiting reward g^* . In particular, they show that W^g is continuous and monotone in g and it has a unique root g^* where $W^{g^*} = 0$. This immediately suggests that any numerically efficient root-finding algorithm, for instance bi-section, can be used to find g^* , provided that it is possible to (numerically) compute W^g for any g by solving (3) and (4). In Section 5 and Section 6 we show that this computation involves solving a simple dynamic programming equation or the numerical evaluation of an integral.

To bound the search space for g^* , we can simply rule out values of g that cannot be optimal. First, we must have $g \geq L^* = \min_x L(x)$, as otherwise the expected average cost of any policy exceeds the reward. Therefore the franchisee would never participate in such a game. Second, we must have $g \leq L^* + K$; as otherwise the franchiser will not be prepared to pay a larger reward than the minimal inventory cost L^* plus an ordering cost K for each occurrence of a demand. Based on these observations, in the remainder of the paper we shall only consider rewards $g \in [L^*, L^* + K]$.

4. Characterization of the parametrized optimal stopping problem

We have thus far provided a framework to approach the optimization problem (1) by embedding it in a class of problems parametrized by a reward rate g , and showed that this framework yields a

stationary policy that minimizes the long-run average cost—provided that the g -revised problem has a solution. In this section, we concentrate on the g -revised problem in isolation and study the structure of its optimal solution.

Based on a *geometric* analysis, we *conjecture* that the solution has a very specific form. This is of particular importance, as the structure of this policy immediately carries over to the original inventory problem under consideration. Because in the remainder the reward g is fixed, we suppress the dependency on g in the notation.

Recall that our first problem is the optimal stopping problem (3)

$$W(x) = K + \inf_{\tau} \mathbb{E}_x \left[\int_0^{\tau} (L(I_s)) - g \, ds \right]$$

with running cost $L(\cdot)$ and terminal cost K . The aim is at finding an optimal stopping time τ that attains the infimum (assuming it exists) for a given initial inventory level x .

From this definition of $W(x)$, we *infer* that the franchisee prefers, on the one hand, to confine the inventory process to states where $L(\cdot) - g < 0$, because there a net profit results. On the other hand, as the inventory drifts to the left and $L(x) - g > 0$ as $x \rightarrow -\infty$, it is eventually optimal to place an order. Based on this intuition, we define the *continuation* and *stopping sets*, respectively, as

$$\mathcal{C} = (\ell, r], \quad \mathcal{D} = (-\infty, \ell], \quad (6a)$$

where

$$\mathcal{A} = \{x : L(x) - g < 0\}, \quad \ell = \inf \mathcal{A}, \quad r = \sup \mathcal{A}. \quad (6b)$$

It is clear that the quasi-convexity of L implies that $L(x) \geq g$ everywhere on \mathcal{D} . Moreover, since the inventory drifts to the left, it must be optimal to stop the process once it enters \mathcal{D} ; the cost cannot decrease when staying any longer in \mathcal{D} . The actual challenge is to show that it is optimal not to stop somewhere in \mathcal{C} . In other words, the problem is to prove that the *optimal stopping time* τ can be uniquely characterized as the hitting time of \mathcal{D} , i.e., as

$$\tau = \inf \{t : I_t \in \mathcal{D}\}. \quad (7)$$

It is important to realize that taking τ as a hitting time of the stopping set \mathcal{D} has crucial implications for the solution of the g -revised problem in (4). First, it suggests that the stopping criterion is not affected by the initial inventory level x . Second, the fact that \mathcal{D} is a half-ray implies that the optimal ordering policy is of threshold type, that is, it is optimal to place an order when the inventory level is smaller than or equal to the threshold ℓ , and it is optimal not to place an order when the process is above this level. Moreover, as we will see, the optimal reorder level S , i.e., the solution of (4), must lie in the continuation set \mathcal{C} . Finally, as these properties are topological, so to say, it is evident that they hold for any g , hence in particular for the long-run average optimal policy π^* associated with g^* which solved $W^{g^*} = 0$. Consequently, the optimal policy π^* is of (s, S) -type.

The above observations suggest to solve the g -revised problem in two consecutive steps. First, prove that the hitting time (7) is an optimal stopping time for (3). Then, find the optimal S in $W = \min_S W(S)$.

5. Periodic-review systems with discrete demand

In this section we demonstrate how the formulation of the inventory control problem as a (g -parametrized) optimal stopping problem can be used to solve the periodic-review inventory problem with discrete demand. In particular we will show that the optimality of (s, S) -policies becomes geometrically evident when the inventory cost function L is quasi-convex. Then we illustrate the procedure by means of an example of Zheng and Federgruen (1991). In Section 5.3 we compare the g -revised problem to the classical formulation in terms of a Markov decision problem (MDP). Finally, in Section 5.4 we use a linear programming approach, rather than Bellman's equation, to obtain the optimality result.

5.1. Proof of optimality with Bellman's Equation

As we deal here with a periodic-review, discrete demand inventory problem, we modify our notation slightly. Let I_n be the start-of-period inventory and Y_n the demand in the n th period. Based on the end-of-period inventory $I_n - Y_n$ we either decide to stop or to continue for at least one more

period. In case of the latter, $I_{n+1} = I_n - Y_n$. For a given stopping policy time τ and reward g , the expected cost is

$$W(i, \tau) = K + \mathbb{E}_i \left[\sum_{n=1}^{\tau} (L(I_n) - g) \right], \quad (8)$$

where $L(i) = \mathbb{E}[h(i - Y)]$ and Y is the one-period demand and i is the initial inventory level. The aim is to find an optimal stopping time, i.e., a rule τ^* such that $W(i, \tau^*) = \inf_{\tau} W(i, \tau)$, and the value function, i.e., the function $v(i) = W(i, \tau^*)$ for general starting point i . As we will see, the stopping time τ^* specifies the reorder level s and with v we can find the optimal order-up-to level S . Note again that v , S , and so on, depend on g ; we suppress this in the notation.

In order to account for the discrete demands we slightly adjust our definitions of the continuation and stopping sets in (6), and we let $\mathcal{C} = \{i : L(i) - g < 0\}$ be the continuation set, $\ell = \min \mathcal{C}$, $r = \max \mathcal{C}$, and $\mathcal{D} = \{i : i < \ell\}$ be the stopping set.

LEMMA 2. *For $i \in \mathcal{D}$ it is optimal to stop immediately.*

Proof. The inventory process cannot go up when no order is placed, hence when $I_n \in \mathcal{D}$ for some n , I will remain in \mathcal{D} until we stop. As on $i \in \mathcal{D}$, the cost $L(i) - g \geq 0$, each such term only increases the cost $W(S, \tau)$. \square

Clearly, it is possible to stop the process right away, which results in paying the terminal cost K . Thus, $v(\cdot)$ must be dominated by K . Moreover, v must be smaller than the one-step cost, for otherwise we will also stop right away. In other words, $v(i) \leq \mathbb{E}[v(i - Y)] + L(i) - g$. Taken together, we see that v must satisfy the two constraints

$$v(i) \leq K, \quad (9a)$$

$$\mathbb{E}[v(i - Y)] - v(i) \geq g - L(i). \quad (9b)$$

The combination of these two inequalities implies the following Lemma.

LEMMA 3. *$v(i) < K$ for $i \in \mathcal{C}$.*

Proof. Clearly, $L(i) - g < 0$ for $i \in \mathcal{C}$. Thus, on \mathcal{C} , $v(i) \leq \mathbb{E}[v(i - Y)] + L(i) - g < \mathbb{E}[v(i - Y)] \leq K$. \square

As we just have two choices at each point i , the value function v must satisfy the dynamic programming equations:

$$v(i) = \min\{K, \mathbb{E}[v(i - Y)] + L(i) - g\}. \quad (10)$$

With this we can construct v from ‘left to right’. By Lemma 2, $v(i) = K$ for $i \in \mathcal{D}$. Thus, the first element of interest is ℓ . Clearly, $\mathbb{E}[v(\ell - Y)] \leq K$ while $L(\ell) - g < 0$. Hence, at ℓ we take $v(\ell) = \mathbb{E}[v(\ell - Y)] + g - L(\ell)$, as the right hand side is less than K . By induction it follows that $v(i) = \mathbb{E}[v(i - Y)] + g - L(i)$ for all $i \in \mathcal{C}$. As a result, the constraint (9b) is binding on \mathcal{C} , which has a nice consequence.

LEMMA 4. $v(i)$ cannot be minimal when $L(i) - g \geq 0$.

Proof. Let v attain a minimum at i . Then, by Lemma 3 and the construction above, $v(i) < K$, hence constraint (9b) is binding at the minimizer i . If we combine this with the assumption that $g - L(i) < 0$ at i , we get

$$\mathbb{E}[v(i - Y)] - v(i) = g - L(i) < 0 \implies \mathbb{E}[v(i - Y)] < v(i).$$

The right hand states that the expectation of a function is smaller than its minimum, but this can never be true. Thus, at any minimizer i it must be that $L(i) - g \leq 0$.

Finally, points where $L(i) = g$ can be excluded, for at such i , $v(i) = \mathbb{E}[v(i - Y)]$. Such i do not affect the solution v . \square

As a consequence of the above reasoning, the franchisee will prefer to rebound (reorder) to the point S at which the value function v takes its minimum. Thus, taken the above lemmas together, we conclude the following.

THEOREM 3. For $i \in \mathcal{D}$, $v(i) = K$, and for $i \in \mathcal{C}$, $v(i) < K$. Hence, on \mathcal{D} it is optimal to stop and on \mathcal{C} it is optimal to continue. Moreover, the optimal starting level S lies in \mathcal{C} .

For quasi-convex L the optimality of (s, S) -policies becomes a simple corollary of the above.

COROLLARY 1. When $L(\cdot)$ is quasi-convex and $\mathbb{E}[L(i - Y)] > K$ for $i \rightarrow \pm\infty$, (s, S) -policies are optimal with $s = \ell - 1$ and S the minimizer of the value function v of the dynamic programming equation (10).

Proof. Since by assumption, $L(\cdot)$ is quasi-convex, we can characterize \mathcal{C} by its left and right limit as $\mathcal{C} = \{\ell, \dots, \mathfrak{r}\}$. At \mathcal{D} we stop, hence, below ℓ we place an order. We order up to S , which lies in \mathcal{C} by Lemma 4. The other condition implies that L is sufficiently large, so it cannot be optimal to let the process drift to $\pm\infty$. \square

REMARK 1. It is important to remark that the dynamic programming equation cannot be extended to models with continuous demand, as the real numbers cannot be enumerated from left to right. To address the continuous case we solve, in Section 6, the related optimal stopping theory with free-boundary theory.

5.2. An illustrative example

Let us show how to apply the above to a numerical example from Zheng and Federgruen (1991). The example is characterized by the following input parameters: $Y \sim \text{Poisson}(15)$, $K = 64$, and $h(i) = 9 \min\{i, 0\} + 1 \max\{i, 0\}$. The optimal policy for this example is an (s, S) -policy with optimal values $s^* = 10$ and $S^* = 49$, and the optimal average cost per period $g^* \approx 42.7$.

In Figure 5.2 (left panel) we show a plot of $L(\cdot) - g$ and the solution v of the dynamic programming equation (10) for a reward $g = 45$. It can be seen that $\ell = 10$ and $\mathfrak{r} = 60$, and a numerical inspection of the graph of v shows that its minimum S lies around $S = 50$, which is definitely smaller than $\mathfrak{r} = 60$, in accordance to Lemma 4. Moreover, the value $g = 45$ is larger than the optimal ≈ 42.8 ; indeed we see that $v(S) < 0$, as it should in accordance with the reasoning that lead to (5). In Figure 5.2 (right panel) we show that if g is updated to 42, the minimum of (the recomputed value function) v is now positive, but still S lies to the left of \mathfrak{r} . Finally, bisecting on g readily yields the optimal values s^* , S^* and g^* .

5.3. Comparison with the classical MDP formulation

Iglehart (1963, Section 4), constructs a set of numbers $\psi = \{\psi(i)\}$ and a constant $g > 0$ that satisfy

$$\psi(i) + g = \min_{j \geq i} \{K \mathbf{1}\{j > i\} + L(j) + \mathbb{E}[\psi(j - Y)]\}. \quad (11)$$

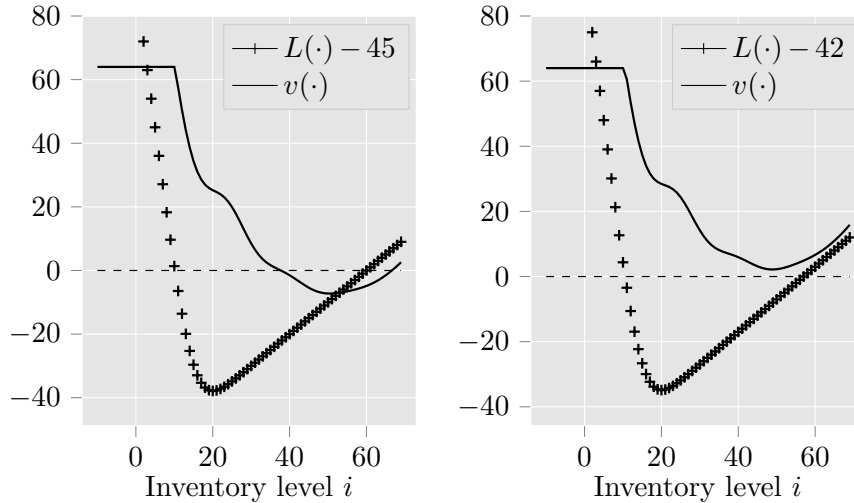


Figure 1 The g -revised inventory cost $L(i) - g$ and the solution v of the related dynamic programming equation (10) for $g = 45$ (left) and $g = 42$ (right).

Let us discuss the relation between this equation and the g -revised approach.

Assuming that π^* is an (s, S) -policy, suppose that π^* prescribes not to order in state i . Then $j = \pi^*(i) = i$, and (11) reduces to

$$\psi(i) = \mathbb{E}[\psi(i - Y)] + L(i) - g$$

which is precisely (9b). If, on the other hand, an order takes place in i , i.e., i lies in the stopping set \mathcal{D} , then $S = \pi^*(i) > i$ as we place an order up to level S . In this case (11) becomes,

$$\psi(i) = K + \mathbb{E}[\psi(S - Y)] + L(S) - g.$$

Clearly, as S lies in the continuation set \mathcal{C} , we can evaluate the previous equation at $i = S$, i.e., $\psi(S) = \mathbb{E}[\psi(S - Y)] + L(S) - g$. Then subtracting these two equations results in the following condition for $i \in \mathcal{D}$:

$$\psi(i) - \psi(S) = K.$$

This becomes equivalent to $v(i) = K$ for $i \in \mathcal{D}$ if we impose the condition $\psi(S) = 0$. Now observe that this condition closes the circle: the choice $\psi(S) = 0$ precisely corresponds to the root $W^{g^*}(S^*) = 0$ we searched for in Section 3.2 to guarantee the existence of an optimal policy for the inventory problem.

Note that the solution ψ of (11) can be used to find an optimal (s, S) -policy. However, since ψ is unbounded, it seems not to be possible to conclude that the associated (s, S) -policy is optimal in the space of stationary policies. To by-pass this point, it is proposed in Zheng (1991) (and in Chen and Feng (2006)) to modify the inventory system such that orders can be returned, and for such modified systems the (modified) ψ is bounded.

When using the g -revised approach such modifications are not necessary, as here, unlike (11), the decision when to order and the decision how much to order are decoupled. The first decision is solved by the dynamic programming equations (10), and this has a unique solution v as \mathcal{D} acts as an obligatory stopping set, c.f., Lemma 2. The second decision, i.e., to determine the optimal starting level S , just uses this value function v , but does not affect its existence in any way. Moreover, note that in (10) there are just two options to choose in the minimization, whereas in (11) the action space contains an infinite number of elements.

5.4. Proof of optimality with linear programming

In this section, we use an idea in Yushkevich and Dynkin (1969, Chapter 3) to formulate the optimal stopping problem (8) as a linear program. Then, by using duality theory, we provide yet another proof of the optimality of (s, S) -policies for the original inventory control problem. Interestingly, this approach is straightforward. We first formulate a primal and a dual linear optimization problem. Then, with complementary slackness arguments we conjecture a solution for the primal and a solution for the dual problems, and show that both solutions are feasible. Finally, by proving that there is no duality gap, that is, the value of the objective of the primal and dual solutions are the same, we can conclude that the conjectured solutions are optimal. As a matter of fact, with this procedure the proof reduces to elementary algebra.

As a first step, to prevent to deal with an infinite dimensional linear programming problem, we assume that the demand is finite and bounded by m . Then, we can reduce the state space of the inventory process to a finite set. The upper bound $r = \max \mathcal{C}$ follows right away from Lemma 4. As a lower bound we can take $M = \ell - 1 - m$ since we know that $L(i) - g \geq 0$ for all $i < \ell$ and

the demand cannot exceed m . With these bounds we henceforth consider a linear program on the space $\mathcal{S} = \{M, \dots, r\}$.

Let v be the vector (v_M, \dots, v_r) , $\mathbf{1} = (1, \dots, 1)$ and I the identity matrix, both of suitable dimensions¹.

Let us next consider the inequalities in (9). The first inequality (9b) can be rewritten as

$$\mathbb{E}[v(i - Y)] - v(i) = (Pv)(i) - v(i) \geq g - L(i), \quad (12)$$

where P is the transition matrix as generated by the demand, i.e., $P_{ij} = \mathbb{P}(Y = i - j)$ when $j > M$ and, to handle the boundary effects, $P_{iM} = \mathbb{P}(Y \geq i - M)$ otherwise. Combining this with the second inequality (9a) we get the following lemma.

LEMMA 5. *The optimal stopping problem (8) is equivalent to the standard linear problem*

$$\max \quad \mathbf{1} \cdot v \quad (13a)$$

$$\text{subject to} \quad (I - P)v \leq L - g \quad (13b)$$

$$v \leq K \quad (13c)$$

LEMMA 6. *The dual of (13) is given by*

$$\min \quad \mu \cdot (L - g) + \mathbf{1} \cdot K \quad (14a)$$

$$\text{subject to} \quad \mu(I - P) \leq \mathbf{1} \quad (14b)$$

$$\mu \geq 0 \quad (14c)$$

We henceforth address the linear programs (13) and (14) as the *primal* and *dual* problems, respectively.

The form of the dual objective is revealing. Because $\mu \geq 0$, it is tempting to guess that the optimal μ is such that $\mu_i > 0$ for $L(i) - g < 0$, and $\mu_i = 0$ otherwise.

By decomposing μ into sub-vectors $\mu_{\mathcal{D}}$ and $\mu_{\mathcal{C}}$, we can write this guess in a more condensed form

as $\mu_{\mathcal{C}} > 0$ and $\mu_{\mathcal{D}} = 0$. With this decomposition of the state space, the transition matrix P can be written in terms of its sub-matrices as

$$P = \begin{pmatrix} P_{\mathcal{D}\mathcal{D}} & P_{\mathcal{D}\mathcal{C}} \\ P_{\mathcal{C}\mathcal{D}} & P_{\mathcal{C}\mathcal{C}} \end{pmatrix}.$$

Because the inventory process is non-increasing and \mathcal{D} lies to the left of \mathcal{C} , P is lower-triangular, in particular, $P_{\mathcal{D}\mathcal{C}} = 0$ and $P_{\mathcal{C}\mathcal{C}}$ is sub-stochastic.

Next, we relate this guess for μ to the constraints of the primal problem. From complementary slackness, we are lead to assume that the following complementary slackness relations hold:

$$\mu_{\mathcal{D}} = 0 \iff [(I - P)v]_{\mathcal{D}} < L_{\mathcal{D}} - g$$

$$\mu_{\mathcal{C}} > 0 \iff [(I - P)v]_{\mathcal{C}} = L_{\mathcal{C}} - g.$$

Thus, the inequality $(I - P)v = L - g$ should be binding on \mathcal{C} . In terms of the primal problem, this means that on \mathcal{C} it is optimal not to stop the process, which in turns implies that $v_{\mathcal{C}}$ should be smaller than K . By similar reasoning from the first relation, it must follow that $v_{\mathcal{D}} = K$. We thus have identified a potential solution for the primal problem.

LEMMA 7. *The vector $v = (v_{\mathcal{D}}, v_{\mathcal{C}})$ such that*

$$v_{\mathcal{D}} = K, \tag{15a}$$

$$v_{\mathcal{C}} = (I - P_{\mathcal{C}\mathcal{C}})^{-1}(L_{\mathcal{C}} - g + P_{\mathcal{C}\mathcal{D}}K) \tag{15b}$$

is a feasible solution for the primal problem.

We now use the above constraints on v to guess a second set of complementary slackness relations:

$$v_{\mathcal{D}} = K \iff [\mu(I - P)]_{\mathcal{D}} < 1$$

$$v_{\mathcal{C}} < 0 \iff [\mu(I - P)]_{\mathcal{C}} = 1$$

Hence, the inequality $\mu(I - P) = 1$ should be binding on \mathcal{C} , and non-binding on \mathcal{D} . In terms of the dual problem, this means that it is optimal not to stop the process on \mathcal{C} , hence $\mu_{\mathcal{C}} > 0$. In a similar vein, it follows that $\mu_{\mathcal{D}} = 0$. We thus found a candidate solution for the dual problem.

LEMMA 8. *The vector $\mu = (\mu_{\mathcal{D}}, \mu_{\mathcal{E}})$ such that*

$$\mu_{\mathcal{D}} = 0, \tag{16a}$$

$$\mu_{\mathcal{E}} = 1(I - P_{\mathcal{E}\mathcal{E}})^{-1} \tag{16b}$$

is a feasible solution for the dual problem.

The above complementary slackness relations guided us to guess the specific form of the solution (15) for the primal problem and (16) for the dual. By Lemma 7 and Lemma 8 these are primal, and dual, feasible. And now, as is often the case with primal-dual arguments, it is irrelevant how we found the solutions: if we can show that the objectives of the primal and the dual problem have the same value with these solutions, both solutions are optimal.

LEMMA 9. *The solutions of Lemma 7 and 8 are optimal as there is no duality gap.*

Clearly, we can again show that Theorem 3 holds, but now with linear programming and duality theory.

6. Continuous-review systems with continuous demand

In this section we will once again be concerned with inventory continuous-review systems processes with continuous demand. In this case the value function of the optimal stopping problem cannot be easily obtained by means of a recursive procedure, such as the dynamic programming equations, as we could in the discrete demand case of Section 5. To deal with this (much) more complicated case we follow the same reasoning, but we use the more extensive tools of free boundary theory to prove that there exists solutions of the associated optimal stopping problems, parametrized by the reward g .

The proof of existence of a solution of the optimal stopping problem proceeds along a number of steps.

By the same reasoning as in Section 5.1, we obtain that the value function V of the optimal stopping problem (3) must satisfy the continuous analogues of the two inequalities (9). Moreover,

it is reasonable to assume that $V(x)$ satisfies at least one of these inequalities at any x . Thus, and similar to Lemma 5, we infer that V must be the largest function that satisfies these inequalities.

These observations lead us to consider the following *free-boundary problem*. Find sets \mathcal{C} and \mathcal{D} that maximize the solution v of the system

$$\mathcal{L}v \geq g - L(x), \tag{17a}$$

$$v \leq K, \quad (v < K \text{ on } \mathcal{C} \text{ and } v = K \text{ on } \mathcal{D}). \tag{17b}$$

Here

$$(\mathcal{L}f)(x) = \lambda(\mathbb{E}[f(x - Y)] - f(x)) \tag{17c}$$

is the infinitesimal operator of the inventory process (I_t) acting on functions f in its domain, which here can be taken as the set of bounded and measurable functions on \mathbb{R} . In the sequel we scale the time (without loss of generality) such that $\lambda = 1$, so that we do not have to carry λ in the notation.

Observe that the set \mathcal{C} and its boundary \mathcal{D} are not fixed, but are part of the problem, hence the name ‘free-boundary problem’. We refer to Peskir and Shiryaev (2006) for further background on the relation between optimal stopping theory and free-boundary problems.

In the sequel we first provide a guess for the optimal sets \mathcal{C} and \mathcal{D} . Then we prove that the system (17) has a solution \hat{v} on $\mathcal{C} \cup \mathcal{D}$. In the last step we prove with verification that $\hat{v} = V$, in other words, we show that our guess for \mathcal{C} and \mathcal{D} is optimal and that the solution \hat{v} of the free-boundary problem (17) also solves the optimal stopping problem (3).

We use the graph of $L(\cdot) - g$, and the quasi-convexity of $L(\cdot)$ to *conjecture* that the sets $\mathcal{C} = (\ell, r]$ and $\mathcal{D} = (-\infty, \ell]$, as defined by (6), are the optimal continuation and stopping sets in the free-boundary problem (17).

Recall from Lemma 4 that in the discrete case, the franchisee will only place an order somewhere in (the discrete analog of) \mathcal{C} . A similar result holds for the continuous case.

LEMMA 10. *A solution v of the system (17) attains its minimum in \mathcal{C} .*

Proof. The reasoning of the proof of Lemma 4 carries over straightaway to the continuous case.

□

As a consequence, we are not interested in the solution of the free-boundary problem beyond \mathcal{C} . Hence, the (in)equalities in problem (17) can be restricted to $(-\infty, r]$.

Using the two previous steps, we reduce problem (17) to finding a solution for the system

$$(\mathcal{L}v)(x) = g - L(x), \quad \text{for } x \in \mathcal{C}, \quad (18a)$$

$$(\mathcal{L}v)(x) = 0, \quad \text{for } x \in \mathcal{D}, \quad (18b)$$

$$v(x) < K, \quad \text{for } x \in \mathcal{C}, \quad (18c)$$

$$v(x) = K, \quad \text{for } x \in \mathcal{D}. \quad (18d)$$

As the next two lemmas show, there exists a unique solution for this system.

LEMMA 11. *The system*

$$(\mathcal{L}v)(x) = g - L(x), \quad x > y, \quad (19a)$$

$$v(x) = K \quad x \leq y, \quad (19b)$$

has a unique solution \hat{v} for any fixed $y \in \mathcal{D} \cup \mathcal{C}$.

LEMMA 12. *The solution \hat{v} of (19) associated with the choice $y = \ell$ solves (18) uniquely.*

It remains to prove that the solution \hat{v} of (19) also solves the optimal stopping problem over all admissible stopping times τ .

THEOREM 4. *The solution \hat{v} of (19) with $s = \ell = \inf \mathcal{C}$ solves the optimal stopping problem (3).*

Thus, the hitting time $\tau = \inf\{t \geq 0 : I_t \in \mathcal{D}\}$ is an optimal stopping time.

7. Discussion

The framework presented in the current paper shows that the problem of finding optimal policies for single-item inventory systems with fixed ordering costs can be reduced to the study of a class of parametrized optimal stopping problems. The literature on optimal stopping and parametrized

optimization have a long history and offer many powerful tools to analyze such problems. Our framework opens up the possibility of approaching inventory control with these tools, as an alternative to developing ingenious concepts specific to inventory theory. An important next step therefore is to apply the proposed framework to further (more general or restricted) inventory control problems besides the class of problems we analyzed in this paper. Below we provide a number of such examples. Finally, we discuss the limitations of our approach.

7.1. Constant demand

As a simple, first, case, consider the standard EOQ model where demand arrives at a constant and continuous rate. To cater for this, take $\mathbb{P}(Y = d/\lambda) = 1$ where d corresponds to demand. Then, the generator of the inventory process becomes $(\mathcal{L}v)(x) = \lambda(v(x - d/\lambda) - v(x))$, hence

$$\lim_{\lambda \rightarrow \infty} (\mathcal{L}v)(x) = -d \frac{d}{dx} v(x).$$

If no backlogging is allowed, we take the hitting set $(-\infty, 0]$ as stopping set, i.e, once the inventory becomes zero (or lower), place an order. Otherwise, choose \mathcal{D} in accordance with (6). A plain substitution in (18) yields the well-known EOQ formulas with or without backlogging.

The existence and optimality proof under the average cost criterion is easy to provide with the approach of Section 6. For other proofs we refer to Beyer and Sethi (1998) and Wee et al. (2009).

7.2. Poisson demand

The direct generalization of the EOQ setting is the case with unit-sized demands that arrive as a Poisson process with rate λ . In this case, $(\mathcal{L}v)(i) = \lambda(v(i - 1) - v(i))$, as the inventory process I is an integer due to the fact that the demands are integers. If we additionally restrict I to the non-negative integers, take the holding cost to be the linear function hi , and assume that the leadtime is zero, then (19) reduces to the system

$$\begin{aligned} v(i) &= \frac{hi - g}{\lambda} + v(i - 1), \quad i = 1, 2, \dots, \\ v(i) &= K, \quad i = 0, -1, -2, \dots \end{aligned}$$

From this, $v(i) = K + hi(i + 1)/2\lambda - gi$.

According to the solution scheme described in Section 3 we first fix g and search for the minimizer S of v to yield the optimal order quantity. By considering the inequalities $v(S-1) \geq v(S)$ and $v(S) \leq v(S+1)$, we obtain that, for given g , S must satisfy,

$$S \leq \frac{g}{h} \leq S+1$$

We next need to find the optimal g . For this, define $v_g = K + hS(S+1)/2\lambda - gS$, where S satisfies the above condition. After a bit of algebra, we finally obtain that the optimal order quantity S must be such that

$$S(S-1) \leq \frac{2\lambda K}{h} \leq S(S+1).$$

This is precisely the result found earlier by Sivazlian (1974).

7.3. Demand processes with constant and stochastic components

In Hordijk and Van der Duyn Schouten (1986) and Presman and Sethi (2006) the demand process is given by $Z_t = dt + \sum_{n=1}^{N_t} Y_n$, where d represents a constant, deterministic component of the demand. Observe that if $d=0$ we have the model of Section 2; if $\lambda=0$ we have the EOQ model with backlogging. Adding a such deterministic component to the demand, requires to change the generator associated with the inventory process from (17c) into

$$(\mathcal{L}f)(x) = -q \frac{d}{dx} f(x) + \lambda(\mathbb{E}[f(x-Y)] - f(x)).$$

The rest of the solution scheme of Section 6 remains in place, save that the analysis becomes more technical. Besides this, with our methods we are able to actually compute an optimal policy, an aspect that is missing in both Hordijk and Van der Duyn Schouten (1986) and Presman and Sethi (2006). As a further extension, we conjecture that with our procedure it is possible to handle Levy demand processes.

7.4. Non-quasi-convex inventory costs

To keep the proofs clean we assume that the inventory cost function $L(\cdot)$ is quasi-convex, but for the numerical analysis this is by no means necessary. It is evident from the dynamic programming equations (10) that very complicated inventory cost functions can be analyzed numerically. (Observing

that the evaluation of integrals such as $\mathbb{E}[L(x - Y)]$ can typically only be carried out numerically, we only consider discrete demand.) In such cases, the continuation set $\mathcal{A} = \{x : L(x) - g < 0\}$ splits into a number of (disjoint) intervals (ℓ_i, r_i) . The rest of the analysis (save the structure results for which the quasi-convexity of L is a sufficient condition) carries over directly to this more general case.

7.5. Joint-ordering

We can analyze multi-item inventory problems quite easily, as long as a fixed ordering cost K is charged on each order as a whole and not on orders of individual items. Let us sketch how to proceed for a two-item inventory process. The demands for each type are given by the sequences of i.i.d. random variables $\{Y_{n,i}\}$, $i = 1, 2$ and arrive according to independent Poisson processes with rates λ_i . The inventory cost needs to be extended to accommodate, in the obvious way, to two item types. Similar to the single-item case, for given g , we need to determine a stopping set \mathcal{D} and a continuation set \mathcal{C} , and solve a free-boundary problem as in (17) but with generator

$$(\mathcal{L}f)(x) = \lambda_1(\mathbb{E}[f(x - Y_1)] - f(x)) + \lambda_2(\mathbb{E}[f(x - Y_2)] - f(x)),$$

The guess for the stopping set is evident: it is the set of inventory levels (x_1, x_2) ‘south-west’ of the boundary of $\mathcal{C} = \{(z_1, z_2) \in \mathbb{R}^2 : L(z_1, z_2) - g < 0\}$. Finally, take the value function equal to K on this set, and solve the boundary problem equivalent to (18). The minimum of the value function must again lie in \mathcal{C} , by the same reasoning as used in Lemma 4.

7.6. Limitations

A key assumption in the framework is that the ordering cost K is fixed, as can be observed in all the examples above. To understand the reason for this, recall that in Section 3.1 we subsequently solve for the optimal stopping time τ , in (7), and then for the optimal ordering quantity, in (4). Because the ordering cost K is independent of the order size, the inventory control problem can be decomposed into these two problems. However, when the order cost is a function of the order size, the analysis becomes more difficult, as now a relation between the termination cost, given by some (positive) function $K(\cdot)$, and the initial inventory level comes into play. We believe that the

concept of parametrization with a reward g will still be helpful in obtaining (geometric) insight into such problems; whether optimal stopping theory can cope with such generalizations we, as yet, do not know.

Appendix A: Proofs of results in Section 3

Proof of Lemma 1. (a) Let $g' \leq g$ and $(S^{g'}, \tau^{g'})$ and (S^g, τ^g) be the optimal policies (which exist by assumption) associated with g' and g such that $W^{g'} = \inf_x \inf_\tau W^{g'}(x, \tau) = W^{g'}(S^{g'}, \tau^{g'})$ and $W^g = \inf_x \inf_\tau W^g(x, \tau) = W^g(S^g, \tau^g)$.

For any (finite) stopping time τ and initial inventory level x , we have

$$\mathbb{E}_x \left[\int_0^\tau (L(I_z) - g') dz \right] \geq \mathbb{E}_x \left[\int_0^\tau (L(I_z) - g) dz \right].$$

Thus, we also have

$$\begin{aligned} W^{g'} - K &= \mathbb{E}_{S^{g'}} \left[\int_0^{\tau^{g'}} (L(I_z) - g') dz \right] \geq \mathbb{E}_{S^{g'}} \left[\int_0^{\tau^{g'}} (L(I_z) - g) dz \right] \\ &\geq \inf_\tau \mathbb{E}_{S^{g'}} \left[\int_0^\tau (L(I_z) - g) dz \right] = W^g(S^{g'}, \tau^g) - K. \end{aligned}$$

Finally, since $W^g = W^g(S^g, \tau^g) = \inf_x W^g(x, \tau^g)$, we conclude that $W^{g'}(S^{g'}, \tau^g) \geq W^g(S^g, \tau^g) = W^g$.

Hence, $W^{g'} \geq W^g$.

(b) Let (S, τ) be an arbitrary policy and suppose g^- is its average cost. Then, by the renewal reward theorem, we have

$$\begin{aligned} 0 &= \mathbb{E}_S \left[\int_0^\tau (L(I_z) - g^-) dz \right] + K \\ &\geq \inf_x \inf_\tau \mathbb{E}_x \left[\int_0^\tau (L(I_z) - g^-) dz \right] + K = W^{g^-}. \end{aligned}$$

That is, $W^{g^-} \leq 0$ if g^- is the average cost of some arbitrary policy. If we let $g^+ = 0$. Then, it is evident that $W^{g^+} = W^0 = \inf_x \inf_\tau \mathbb{E}_x \left[\int_0^\tau L(I_z) dz \right] + K \geq 0$.

(c) We first prove the continuity for a given policy (x, τ) . By Lemma 3 we can split the expectation in the definition of $W^g(\cdot, \cdot)$ as

$$W^g(x, \tau) = \mathbb{E}_x \left[\int_0^\tau L(I_z) dz \right] - g \mathbb{E}_x[\tau] + K$$

which is obviously continuous in g .

Next, we prove that $g \rightarrow W^g(x) = \inf_\tau W^g(x, \tau)$ is continuous for fixed x . This is not immediate, as the infimum here is taken over an uncountable set of stopping times. Moreover, the stopping times τ^g and $\tau^{g'}$ that attain the infima need not be the same for different rewards g and g' . To get around this problem, assume $g = g' + \phi$ for some $\phi > 0$. Because $W^g(x)$ is an infimum, there exists for any $\epsilon > 0$ a stopping time σ such that

$$\begin{aligned}
W^g(x) &\geq \mathbb{E}_x \left[\int_0^\sigma L(I_z) dz \right] + K - g\mathbb{E}_x[\sigma] - \epsilon \\
&= \mathbb{E}_x \left[\int_0^\sigma L(I_z) dz \right] + K - g'\mathbb{E}_x[\sigma] - \phi\mathbb{E}_x[\sigma] - \epsilon \\
&\geq \inf_\tau \mathbb{E}_x \left[\int_0^\tau (L(I_z) - g') dz \right] + K - \phi\mathbb{E}_x[\sigma] - \epsilon \\
&= W^{g'}(x) - \phi\mathbb{E}_x[\sigma] - \epsilon.
\end{aligned}$$

From (a), we have that $W^g(x)$ is decreasing in g . Therefore, it follows that

$$W^{g'}(x) \geq W^g(x) \geq W^{g'}(x) - \phi\mathbb{E}_x[\sigma] - \epsilon. \quad (20)$$

Because ϵ is arbitrary and $\mathbb{E}_x[\sigma] < \infty$ for any admissible stopping time, we have $W^g(x) \rightarrow W^{g'}(x)$ as $\phi \rightarrow 0$.

Finally, we show that $W^g = \inf_x W^g(x) \rightarrow W^{g'} = \inf_x W^{g'}(x)$ as $\phi \rightarrow 0$. This follows from (20) once we can show that $\mathbb{E}_x[\sigma]$ is finite for all x of interest. From Lemma 10, we have that the minimizer S^g of $W^g(\cdot)$ lies in the finite set $\mathcal{A}^g = \{y : L(y) - g \leq 0\}$. Thus, by taking x equal to the supremum r^g of this set, $\mathbb{E}_x[\sigma]$ attains its largest possible value $\mathbb{E}_{r^g}[\sigma]$. Because this is finite, it follows from (20) that

$$\inf_x W^{g'}(x) \geq \inf_x W^g(x) \geq \inf_x W^{g'}(x) - \phi\mathbb{E}_{r^g}[\sigma] - \epsilon,$$

from which we conclude that $g \rightarrow W^g$ is continuous.

□

Appendix B: Proofs of results in Section 5

Proof of Lemma 5. We already derived that the value function v of the stopping problem (8) satisfies conditions (13b) and (13c). It remains to prove the solution of the objective (13a) and the value function v coincide. To establish this, we show that v is the largest function that satisfies constraints (13b)–(13c).

Let w be any solution that satisfies conditions (13b)–(13c), and let τ be any stopping time with finite expectation. Then, using (13b), we obtain

$$\mathbb{E}_i \left[\sum_{n=1}^{\tau} ((I - P)w)(I_n) \right] \leq \mathbb{E}_i \left[\sum_{n=1}^{\tau} (L(I_n) - g) \right].$$

The left hand side is a telescoping sum and equals $w(i) - \mathbb{E}_i[w(I_{\tau+1})]$. Therefore,

$$\begin{aligned}
w(i) &\leq \mathbb{E}_i \left[\sum_{n=1}^{\tau} (L(I_n) - g) \right] + \mathbb{E}_i[w(I_{\tau+1})] \\
&\leq \mathbb{E}_i \left[\sum_{n=1}^{\tau} (L(I_n) - g) \right] + K,
\end{aligned}$$

where we use (13c) to see that $\mathbb{E}_i [w(I_{\tau+1})] \leq K$. Observing that this inequality holds for any stopping time, we conclude

$$w(i) \leq K + \inf_{\tau} \mathbb{E}_i \left[\sum_{n=1}^{\tau} (L(I_n) - g) \right] = v(i).$$

Clearly, then, the value function v must be such that $v(i) \geq w(i)$ for each i . This can be recast into the objective $\max \sum_{i \in \mathcal{S}} v(i)$, which is equal to $\max 1 \cdot v$. \square

Proof of Lemma 6. By means of standard duality arguments, we write the dual problem with dual variables μ as

$$\begin{aligned} \min \quad & \lambda K + \mu(L - g) \\ \text{subject to} \quad & \lambda + \mu(I - P) = 1 \\ & \lambda, \mu \geq 0 \end{aligned}$$

Focusing on the λK term in objective and combining this with the first constraint, we have $\lambda K = [1 - \mu(I - P)]K = 1 \cdot K + \mu(I - P)K = 1 \cdot K$. Thus, the objective of the dual reduces to

$$1 \cdot K + \mu(L - g).$$

This result can be anticipated because the boundary cost K is fixed and cannot be avoided under any policy. Thus, it should not be part of the optimization problem.

Let us observe that λ does not appear in the objective function. Because $\lambda \geq 0$, the constraint $\lambda + \mu(I - P) = 1$ can be expressed as $\mu(I - P) \leq 1$ and λ can be discarded from the linear program altogether. \square

Proof of Lemma 7. Let us first show that v as given by (15) satisfies inequality (13b). For this, write

$$\begin{aligned} L - g &\geq (I - P)v = \begin{pmatrix} I - P_{\mathcal{D}\mathcal{D}} & 0 \\ -P_{\mathcal{C}\mathcal{D}} & I - P_{\mathcal{C}\mathcal{C}} \end{pmatrix} \begin{pmatrix} v_{\mathcal{D}} \\ v_{\mathcal{C}} \end{pmatrix} \\ &= \begin{pmatrix} (I - P_{\mathcal{D}\mathcal{D}})v_{\mathcal{D}} \\ -P_{\mathcal{C}\mathcal{D}}v_{\mathcal{D}} + (I - P_{\mathcal{C}\mathcal{C}})v_{\mathcal{C}} \end{pmatrix}. \end{aligned}$$

This is satisfied on \mathcal{D} because

$$L_{\mathcal{D}} - g \geq (I - P_{\mathcal{D}\mathcal{D}})v_{\mathcal{D}} \geq (I - P_{\mathcal{D}\mathcal{D}})K = 0$$

where the last equality follows as $P_{\mathcal{D}\mathcal{D}}$ is stochastic, therefore $P_{\mathcal{D}\mathcal{D}}K = K$. On \mathcal{C} , substitute (15) into

$$L_{\mathcal{C}} - g = -P_{\mathcal{C}\mathcal{D}}v_{\mathcal{D}} + (I - P_{\mathcal{C}\mathcal{C}})v_{\mathcal{C}} \tag{21}$$

shows that the inequality is true.

Now consider inequality (13c). This is trivially met on \mathcal{D} as $v_{\mathcal{D}} = K$. Now consider the other inequality $v \leq K$. Again, on \mathcal{D} this is trivially met. On \mathcal{C} , let us write $v_{\mathcal{C}} = K - i$; then the aim is to prove that $i \geq 0$. Substituting $v_{\mathcal{C}} = K - i$ into (21) and using that $P_{\mathcal{C}\mathcal{D}}K - P_{\mathcal{C}\mathcal{C}}K = 0$, leads to

$$(I - P_{\mathcal{C}\mathcal{C}})i = g - L_{\mathcal{C}} \geq 0.$$

By Seneta (1981, Theorem 2.1) and the sub-stochasticity of $P_{\mathcal{C}\mathcal{C}}$ it follows that $i \geq 0$, hence $v_{\mathcal{C}} \leq K$. \square

Proof of Lemma 8. Let us first show that μ satisfies inequality (14b). To that end, we write

$$1 \geq \mu(I - P) = (\mu_{\mathcal{D}}, \mu_{\mathcal{C}}) \begin{pmatrix} I - P_{\mathcal{D}\mathcal{D}} & 0 \\ -P_{\mathcal{C}\mathcal{D}} & I - P_{\mathcal{C}\mathcal{C}} \end{pmatrix}.$$

On \mathcal{D} this becomes $1 \geq \mu_{\mathcal{D}}(I - P_{\mathcal{D}\mathcal{D}}) - \mu_{\mathcal{C}}P_{\mathcal{C}\mathcal{D}}$ and on \mathcal{C} we have $1 \geq \mu_{\mathcal{C}}(I - P_{\mathcal{C}\mathcal{C}})$.

By Seneta (1981, Theorem 2.1) and the fact that $P_{\mathcal{C}\mathcal{C}}$ is sub-stochastic, $\mu_{\mathcal{C}} = 1(1 - P_{\mathcal{C}\mathcal{C}})^{-1} \geq 0$. With this result, and as $\mu_{\mathcal{D}} = 0$ by assumption, we see that

$$1 \geq \mu_{\mathcal{D}}(I - P_{\mathcal{D}\mathcal{D}}) - \mu_{\mathcal{C}}P_{\mathcal{C}\mathcal{D}}$$

is satisfied. Hence, (16) satisfies the conditions (14) of the dual problem. \square

Proof of Lemma 9. It is sufficient to show that the respective objective values of v and μ in primal and dual problems are the same. That is,

$$1 \cdot v = 1 \cdot K + \mu(L - g).$$

For the proposed solutions v and μ , we can re-write the left hand side and the right hand side of the equality as

$$\begin{aligned} 1 \cdot v &= 1 \cdot v_{\mathcal{D}} + 1 \cdot v_{\mathcal{C}} \\ &= 1_{\mathcal{D}} \cdot K + 1_{\mathcal{C}} \cdot (I - P_{\mathcal{C}\mathcal{C}})^{-1}(L_{\mathcal{C}} - g + P_{\mathcal{C}\mathcal{D}}K) \\ &= 1_{\mathcal{D}} \cdot K + 1_{\mathcal{C}} \cdot (I - P_{\mathcal{C}\mathcal{C}})^{-1}(L_{\mathcal{C}} - g) + 1_{\mathcal{C}} \cdot (I - P_{\mathcal{C}\mathcal{C}})^{-1}P_{\mathcal{C}\mathcal{D}}K \end{aligned}$$

and

$$\begin{aligned} 1 \cdot K + \mu(L - g) &= 1 \cdot K + \mu_{\mathcal{D}}(L_{\mathcal{D}} - g) + \mu_{\mathcal{C}}(L_{\mathcal{C}} - g) \\ &= 1_{\mathcal{D}} \cdot K + 1_{\mathcal{C}} \cdot (I - P_{\mathcal{C}\mathcal{C}})^{-1}(L_{\mathcal{C}} - g) + 1_{\mathcal{C}} \cdot K. \end{aligned}$$

respectively. Hence, it is sufficient to show that $(I - P_{\mathcal{C}\mathcal{C}})^{-1}P_{\mathcal{C}\mathcal{D}}K = K$. This is true, as

$$P_{\mathcal{C}\mathcal{D}}K = (I - P_{\mathcal{C}\mathcal{C}})K \iff (P_{\mathcal{C}\mathcal{C}} + P_{\mathcal{C}\mathcal{D}})K = K,$$

and $P_{\mathcal{C}\mathcal{C}} + P_{\mathcal{C}\mathcal{D}}$ is a stochastic matrix. \square

Appendix C: Proofs of results in Section 6

Proof of Lemma 11. Use (17c) to rewrite system (19) for $x \geq y$ into the renewal equation

$$\begin{aligned} v(x) &= L(x) - g + \mathbb{E}[v(x - Y)] \\ &= L(x) - g + \int_0^\infty v(x - z) dF(z) \\ &= L(x) - g + K\bar{F}(x - \ell) + \int_0^{x-\ell} v(x - z) dF(z). \end{aligned}$$

Since $L(x)$ and $\bar{F}(x)$ are bounded for $x \in [y, \mathfrak{r}]$, it follows from Karlin and Taylor (1975, Theorem 5.4.1), that the system (19) has a unique and bounded solution \hat{v} . \square

Proof of Lemma 12. We start with checking that \hat{v} satisfies the conditions in (18) for $y = \ell$. First, (19a) is the same as (18a) when $y = \ell$. Next, since $\hat{v} = K$ at the left of $y = \ell$, $(\mathcal{L}\hat{v})(x) = \mathbb{E}[\hat{v}(x - Y)] - \hat{v}(x) = K - K = 0$ for $x \leq \ell$. Thus, \hat{v} also meets conditions (18b) and (18d). To see that (18c) holds, substitute $v(x) = w(x) + K$ into the renewal equation of Lemma 11. Then w must solve

$$w(x) = L(x) - g + \int_0^{x-\ell} w(x - z) dF(z),$$

and $w(x) = 0$ for $x \leq \ell$. Now, again by (Karlin and Taylor 1975, Theorem 5.4.1),

$$w(x) = L(x) - g + \int_0^{x-\ell} (L(x - z) - g) dM(z),$$

where $M(x) = \sum_{i=1}^\infty F_k(x)$ is the renewal function with F_k the k th convolution of F . Clearly, $M(\cdot) \geq 0$ and is non-decreasing. Thus, we have $dM \geq 0$. Because $L(\cdot) - g < 0$ for $x \in \mathcal{C}$, we also have $(L(z) - g) dM(z) \leq 0$. Therefore, $w < 0$ and $v = w + K < K$ on C .

To see that \hat{v} solves (18) only when $y = \ell$, suppose first that $y < \ell$ so that $L(y) > g$. Then

$$\begin{aligned} \hat{v}(y) &= \mathbb{E}[\hat{v}(y - Y)] + L(y) - g && \text{by (19a)} \\ &> \mathbb{E}[\hat{v}(y - Y)] && \text{as } L(y) > g \\ &= K && \text{by (18d)}. \end{aligned}$$

But, by (19b), $v(y) = K$. This rules out that $y < \ell$. On the other hand, if $y > \ell$, $\hat{v}(y) = K$ by (19b), but $v(y)$ must be smaller than K by (18c). \square

Proof of Theorem 4. The proof comes down to a verification theorem, c.f. Peskir and Shiryaev (2006, Theorem 22.1). We first prove that the value function V of (3) is bounded from below by \hat{v} , i.e., $V \geq \hat{v}$.

Then we show that \hat{v} is equal to $K + \mathbb{E}_i \left[\int_0^\tau (h(I_s) - g) ds \right]$ with τ the hitting time of \mathcal{D} . Combining these two arguments implies that \hat{v} attains the infimum in (3), because

$$\begin{aligned} \hat{v}(x) &= K + \mathbb{E}_x \left[\int_0^\tau (h(I_z) - g) dz \right] \\ &\geq K + \inf_\tau \mathbb{E}_x \left[\int_0^\tau (h(I_z) - g) dz \right] = V_x \geq \hat{v}(x), \end{aligned}$$

hence $V = \hat{v}$.

For the first step, note that the process $\{Q_t\}$ defined at time t as

$$Q_t = \hat{v}(I_t) - \hat{v}(x) - \int_0^t (\mathcal{L}\hat{v})(I_z) dz \quad (22)$$

is a martingale with respect to the filtration generated by the demand process (Z_t) , c.f., (Varadhan 2007, Theorem 4.2).

Now \hat{v} satisfies (18) so $\hat{v} \leq K$, hence from (22)

$$K \geq \hat{v}(I_t) = Q_t + \hat{v}(x) + \int_0^t (\mathcal{L}\hat{v})(I_z) dz.$$

When $x \in \mathcal{C}$, $(\mathcal{L}\hat{v}) = g - L$, while for $x \in \mathcal{D}$, $(\mathcal{L}\hat{v}) = 0 \geq g - L(x)$. Thus,

$$K \geq Q_t + \hat{v}(x) + \int_0^t (g - L(I_z)) dz, \quad (23)$$

from which we get that

$$K + \int_0^t (L(I_z) - g) dz \geq \hat{v}(x) + Q_t.$$

Since $\{Q_t\}$ is a martingale, $\mathbb{E}[Q_t] = \mathbb{E}[Q_0] = 0$ as $Q_0 = 0$. Moreover, $\mathbb{E}_x[\tau] < \infty$ for all admissible stopping times, because eventually I_t becomes so small that placing an order becomes cheaper than paying the inventory cost. Therefore, by the optional stopping theorem, $\mathbb{E}[Q_\tau] = 0$. This allows us to replace t by τ in the above inequality and then take the expectation \mathbb{E}_x to conclude

$$K + \mathbb{E}_x \left[\int_0^\tau (L(I_z) - g) dz \right] \geq \hat{v}(x) + \mathbb{E}_x[Q_\tau] = \hat{v}(x). \quad (24)$$

As this holds for any stopping time, it follows that

$$V(x) = K + \inf_\tau \mathbb{E}_x \left[\int_0^\tau (L(I_z) - g) dz \right] \geq \hat{v}(x).$$

For the second step, observe that the stopping τ attains equality in (23). In fact, by (19), $(\mathcal{L}\hat{v})(I_t) = g - L(I_t)$ for $t < \tau$, but when $t = \tau$, $I_\tau \in D$ so that $\hat{v}(I_\tau) = K$. Thus,

$$K = \hat{v}(I_\tau) = Q_\tau + \hat{v}(x) + \int_0^\tau (g - L(I_z)) dz.$$

Finally, taking the expectation \mathbb{E}_x at both sides and again noting that $\mathbb{E}_x[Q_\tau] = 0$, we see that τ also attains the infimum in (24). \square

References

- Axsäter S (2006) *Inventory Control* (Springer), 2nd edition.
- Bather J (1966) A continuous time inventory model. *Journal of Applied Probability* 3(2):538–549.
- Beckmann M (1961) An inventory model for arbitrary interval and quantity distributions of demand. *Management Science* 8(1):35–57.
- Beyer D, Cheng F, Sethi S, Taksar M (2010) *Markovian Demand Inventory Models* (Springer).
- Beyer D, Sethi S (1998) A proof of the EOQ formula using quasi-variational inequalities. *International Journal of Systems Science* 29(11):1295–1299.
- Chen F, Feng Y (2006) Optimization and optimality of (s, S) stochastic inventory systems with non-quasiconvex costs. *Probability in the Engineering and Informational Sciences* 20(02).
- Feng Y, Xiao B (2000) A new algorithm for computing optimal (s, S) policies in a stochastic single item/location inventory system. *IIE Transactions* 32:1081–1090.
- Hadley G, Whitin T (1963) *Analysis of Inventory Systems* (Prentice-Hall, Englewood Cliffs, N.J.).
- Hordijk A, Van der Duyn Schouten F (1986) On the optimality of (s, S) policies in continuous review inventory models. *SIAM J. Appl. Math.* 46(5):912–929.
- Huh W, Janakiraman G, Nagarajan M (2011) Average cost single-stage inventory models: An analysis using a vanishing discount approach. *Operations Research* 59(1):143–155.
- Iglehart D (1963) Dynamic programming and stationary analyses of inventory problems. Scarf H, Gilford D, Shelly M, eds., *Multistage Inventory Models and Techniques*, 1–31 (Stanford, California: Stanford University Press).
- Karlin S, Taylor H, eds. (1975) *A First Course in Stochastic Processes* (Academic Press), 2nd edition.
- Peskir G, Shiryaev A (2006) *Optimal Stopping Problems and Free-Boundary Problems* (Basel: Birkhäuser Verlag).
- Presman E, Sethi S (2006) Inventory models with continuous and Poisson demands and discounted and average costs. *Production and Operations Management* 15(2):279–293.
- Rockafellar R (1974) Conjugate duality and optimization. Technical Report 16, SIAM Publications.
- Scarf H (1959) The optimality of (S, s) policies in the dynamic inventory problem. Arrow K, Karlin S,

- Scarf H, eds., *Mathematical Methods in the Social Sciences*, chapter 13, 196–202 (Stanford: Stanford University Press).
- Seneta E (1981) *Non-Negative Matrices and Markov Chains* (Springer).
- Shiryaev A (2007) *Optimal Stopping Rules* (Springer-Verlag New York, Inc.).
- Sivazlian BD (1974) A continuous-review (s, S) inventory system with arbitrary interarrival distribution between unit demand. *Operations Research* 22(1):65–71.
- Van Foreest N, Wijngaard J (2014) On optimal policies for production-inventory systems with compound Poisson demand and setup costs. *Mathematics of Operations Research* 39(2):533–560.
- Varadhan S (2007) *Stochastic Processes* (Courant Institute of Math. Sciences).
- Veinott A, Wagner H (1965) Computing optimal (s, S) inventory policies. *Management Science* 11(5):525–552.
- Wee H, Wang W, Chung C (2009) A modified method to compute economic order quantities without derivatives by cost-difference comparisons. *European Journal of Operational Research* 194(1):336–338.
- Yushkevich A, Dynkin E (1969) *Markov Processes: Theorems and Problems* (Plenum Press).
- Zabel E (1962) A note on the optimality of (S, s) policies in inventory theory. *Management Science* 9(1):123–125.
- Zheng Y (1991) A simple proof of the optimality of (s, S) policies in infinite-horizon inventory systems. *J. Appl. Prob.* 28:802–810.
- Zheng Y, Federgruen A (1991) Finding optimal (s, S) policies is about as simple as evaluating a single policy. *Operations Research* 39(4):654–665.
- Zipkin P (2000) *Foundations of Inventory Management* (McGraw-Hill/Irwin).