An intuitive approach to inventory control with optimal stopping

Van Foreest, Nicky D.; Kilic, Onur A.

Published in:
European Journal of Operational Research

DOI:
10.1016/j.ejor.2023.05.035

European Journal of Operational Research 311 (2023) 921–924

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Production, Manufacturing, Transportation and Logistics

An intuitive approach to inventory control with optimal stopping

Nicky D. Van Foreest*, Onur A. Kilic

Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands

A R T I C L E I N F O

Article history:
Received 23 November 2022
Accepted 29 May 2023
Available online 30 May 2023

Keywords:
Inventory
Optimal stopping
Dynamic programming

A B S T R A C T

In this research note, we show that a simple application of Breiman's work on optimal stopping in 1964 leads to an elementary proof that (s, S) policies minimize the long-run average cost for periodic-review inventory control problems. The method of proof is appealing as it only depends on the fundamental concepts of renewal-reward processes, optimal stopping, dynamic programming, and root-finding. Moreover, it leads to an efficient algorithm to compute the optimal policy parameters. If Breiman's paper would have received the attention it deserved, computational methods dealing with (s, S)-policies would have been found about three decades earlier than the famous algorithm of Zheng and Federgruen (1991).

1. Introduction

The optimality of (s, S) policies for the single-item inventory model is one of the classical results in the inventory control literature. It is discussed in virtually any book on inventory theory. Scarf (1959) provided a first optimality proof. This work has later been extended to a variety of inventory systems, e.g., finite or infinite planning horizons and long-run average cost or discounted cost objectives. However, for about 30 years there was no efficient procedure available to compute the optimal policy parameters. The first relatively simple optimality proof and construction were given by Zheng (1991) and Zheng & Federgruen (1991). We refer the reader to Beyer et al. (2010) for an interesting and detailed discussion on the subsequent developments.

The aim of the current research note is to show that a simple proof and construction of the long-run average optimal (s, S) policy could have been found in the 1960s if Breiman's (1964) ideas on optimal stopping were applied to inventory control. The approach to the proof and construction is particularly elegant. It does not depend on ingenious (but somewhat specific) concepts such as of K-convexity. Instead, it builds on the fundamental concepts of embedding an optimization problem into a one-parameter family of problems, renewal-reward processes, optimal stopping, dynamic programming, and root-finding. Hence, it has the potential to become a standard approach for inventory textbooks.

The paper is organized as follows. In Section 2, we introduce the inventory problem and model. In Section 3, we provide a brief overview of Breiman's study. In Section 4, we provide the optimality proof and an efficient algorithm to compute the optimal policy. In Section 5, we discuss how the presented approach relates to the existing literature and establish directions for further research.

2. Inventory problem and model

We consider a periodic-review inventory system with backlogging. The demands per period $\{Y_n\}$ form a sequence of independent random variables each distributed as the common non-negative and integer-valued random variable Y. We let Y be the lead time demand if there is a (constant) replenishment lead time. The probability mass of Y is given by $p_j = \mathbb{P}(Y = j)$. We write $(Pf)(i) = \mathbb{E}[f(i - Y)] = \sum_{j\geq 0} p_j f(i - j)$.

Let I_n be the (post-replenishment) inventory level at the start of period n. The inventory process $I = \{I_n\}$ is controlled by an inventory policy π that decides whether to order or not, and if so, the order quantity. In case no replenishment is placed at period n, the inventory level in the next period is $I_{n+1} = I_n - Y_n$, otherwise, it is $I_{n+1} > I_n - Y_n$. We say that a new replenishment cycle starts every period in which a replenishment occurs. Like this, the inventory system progresses from one cycle to the next under policy π.

We let $c(\cdot)$ be the inventory (holding and shortage) cost function. The system incurs an (end-of period) expected cost $L(i) = (Pc)(i)$ when the inventory level is i at the start of the period. We assume that $c(\cdot)$ is such that $L(i)$ is quasi-convex and becomes sufficiently large when $i \to \pm\infty$ to ensure 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. Besides the inventory cost, the system incurs a fixed replenishment cost $K \geq 0$ for each replenishment.

* Corresponding author.
E-mail address: n.d.van.foreest@rug.nl (N.D. Van Foreest).

https://doi.org/10.1016/j.ejor.2023.05.035

We define the long-run average cost under policy π as

$$V_\pi = \limsup_{m\to\infty} \mathbb{E}_\pi \left[\frac{1}{m}\sum_{n=1}^{m} (c(I_n - Y_n) + K\,\mathbf{1}\{I_n > I_{n-1} - Y_{n-1}\}) \right]$$

where $\mathbf{1}\{\cdot\}$ is the indicator function, I_0 is the (known) starting inventory level and $Y_0 := 0$.

In the remainder of the paper we will be concerned with three objectives. The first is to show that there is a policy that achieves the minimal long-run average cost, in other words, there exists a policy π^* with long-run average cost

$$V_{\pi^*} = \inf_\pi V_\pi. \tag{1}$$

The second is to prove that the optimal policy π^* has an (s, S) structure. That is, it is optimal to start each cycle at the same inventory level S and only place an order when the inventory becomes less than or equal to s. The third is to devise an efficient procedure to compute the optimal policy parameters s and S.

3. An overview of Breiman's study

In our analysis, we use several ideas of Breiman (1964) to achieve the objectives mentioned above. But before doing so, it seems fitting to provide a brief overview of the relevant parts of Breiman's study. Sections 10.1–10.4 introduce the concept of an optimal stopping problem. Section 10.5 explains that any stopping rule for a Markov chain can be characterized by two disjoint subsets of the state space: a *stopping* set D and a *continuation* set C. Evidently, the rule tells us to continue when the chain is in C and to stop when it hits D. Section 10.7 uses dynamic programming to find the *value function* $i \to V(i)$ which provides us with the expected payoff starting in some state i until stopping, under the optimal rule. Once V is known, the optimal stopping set can be identified as the set of states i in which $V(i)$ is equal to the reward of stopping directly in state i. The problem of finding the optimal stopping set can be greatly simplified if there exists a set of unfavorable states in which a penalty must be paid to continue and the chain can never leave this set by continuing. Section 10.9 defines this as an *entrance* fee problem. Now, stopping rules cannot be immediately applied to repetitive problems where the system can return to the *origin* and start anew at the expense of a fee. To handle such cases, Section 10.13 discusses *renewal rules* and Section 10.14 explains how to use renewal reward theory to reduce finding an optimal renewal rule for a repetitive problem to finding an optimal stopping rule for just one cycle. The work culminates in Theorem 10.5 which shows that optimal control rules for renewal problems can be found by embedding such problems in a one-parameter family of stopping-rule problems.

For the purposes of our study, we apply the aforementioned ideas in the reverse order. That is, we initially consider the problem of minimizing the total cost over a single replenishment cycle and embed it into a one-parameter family of optimal stopping problems. Then, we reduce this problem to an entrance-fee problem and characterize the associated optimal policy. Next, we show there exists a parameter value for which this optimal policy minimizes the long-run average cost when applied repetitively over consecutive replenishment cycles. This suffices to prove that the long-run average cost optimal policy is an (s, S) policy. Finally, we develop an efficient numerical method to compute the optimal policy parameters. In the following, we provide the details of this sketch and provide references to specific sections of Breiman (1964).

4. Proof and construction of an optimal policy

Suppose that we receive a reward $g > 0$ per period to cover the cost we make to operate the inventory system. In this case the expected inventory cost per period is $L(\cdot) - g$ rather than $L(\cdot)$, so that when $L(i) - g < 0$ we make a net profit. Now consider a cycle in which the inventory level starts at i and we have a rule τ that tell us when to stop the cycle. We can write the expected cost over this cycle as $v(i, \tau) = K + \mathbb{E}_i\left[\sum_{n=1}^{\tau}(L(I_n) - g)\right]$.

We now consider the problem of finding the stopping time that minimizes the expected cycle cost for a given starting inventory level. Following Breiman (1964, Section 10.5), we can formulate this problem as an *optimal stopping problem* in which the goal is to compute the *value function*

$$v(i) = \inf_\tau v(i, \tau) = K + \inf_\tau \mathbb{E}_i\left[\sum_{n=1}^{\tau}(L(I_n) - g)\right]. \tag{2}$$

To solve this optimal stopping problem we introduce two crucially important sets, namely the stopping and continuation sets. These are respectively defined as

$$D = \{i : i \le s\} \quad \text{and} \quad C = \{i : L(i) - g < 0\} \tag{3}$$

where

$$s = \min C - 1. \tag{4}$$

Notice that the optimal stopping problem is trivial if C is empty. Then it is optimal to stop right away at any inventory level, and, we have $v(\cdot) = K$. Therefore we assume henceforth that g is sufficiently large so that C is non-empty. It is clear from its definition that C is closed and bounded as the expected cost function $L(i)$ is quasi-convex and becomes sufficiently large when $i \to \pm\infty$. This suggests that s is well-defined and D is non-empty.

We now characterize the optimal stopping rule for starting inventory levels in D and C. It is evident that it is optimal to stop in D, as it immediately follows by the definition of D and the quasi-convexity of L that $L(i) - g > 0$ for all $i \in D$. Hence, the inventory process can never escape from D without issuing a replenishment order. The following lemma is directly based on Breiman (1964, Section 10.9) and establishes the optimal stopping rule for inventory levels in C.

Lemma 4.1. *The optimal stopping rule τ that solves (2) for $i \in C$ is*

$$\tau = \inf\{n : I_n \in D\}. \tag{5}$$

The results presented above are critical as they show that the optimal stopping rule is the same for all starting inventory levels in D and C. That is, it is optimal to stop when the inventory process hits D. We remark that we have not yet considered the optimal stopping time for inventory levels that are neither in D or C. It will be clear later on that such inventory levels are not relevant for the analysis of the optimal inventory control policy.

Having established the optimal stopping time that minimizes the expected cycle cost, we now turn our attention to the optimal starting inventory level. The value function v gives the expected cycle cost. Hence, we can minimize the expected cycle cost by initiating the cycle at an inventory level where v attains its minimum.

It follows from Breiman (1964, Section 10.7) that the value function v can alternatively be expressed as the solution of the dynamic programming equation

$$v(i) = \min\{K, L(i) - g + (Pv)(i)\}. \tag{6}$$

This equation can easily be solved from left to right, provided the stopping and continuation sets. That is, on D we take $v(i) = K$ and on C we use $v(i) = L(i) - g + (Pv)(i)$ where $(Pv)(i)$ only depends on $v(\cdot)$ to its left as demand is non-negative.

The next lemma uses the above characterization of the value function and provides an upper bound for its minimizer.

Lemma 4.2. *There is a minimizer S of v that lies in C.*

Proof. We establish the proof by showing that $L(S) - g < 0$, which implies that $S \in C$. Let us first observe that $v(i) = K$ on D and $v(i) < K$ on C. The former is evident. The latter follows from $L(i) - g < 0$ which suggests $v(i) \leq L(i) - g + (Pv)(i) < (Pv)(i) \leq PK = K$. Thus we immediately have that the minimizer S cannot be in D and it should satisfy $v(S) < K$. Then, it follows from (6) that $v(S) = L(S) - g + (Pv)(S)$. We proceed by contradiction. Suppose $L(S) - g > 0$. Then we must have $v(S) = L(S) - g + (Pv)(S) > (Pv)(S)$. But this cannot be true as the expectation of a function cannot be strictly lower than its minimum. Finally, suppose $L(S) - g = 0$. Then we must have $v(S) = L(S) - g + (Pv)(S) = (Pv)(S) = \sum_{j \geq 0} p_j v(S-j)$. There are two possibilities. First, we may have $\min_{j\geq 0}\{v(S-j)\} < v(S) < \max_{j \geq 0}\{v(S-j)\}$. This cannot be true as it suggests that the minimum of $v(\cdot)$ must lie to the left of S. Second, we may have $v(S) = v(S-j)$ for all $j \geq 0$ with $p_j > 0$. This cannot be true if there is a $j \geq 0$ with $p_j > 0$ such that $S - j \in C$. Otherwise, we can apply the same reasoning to $S - j$. This completes the proof. □

The policy that solves the optimal stopping problem is also an applicable policy for the original problem (1). The next lemma sheds light onto the average cycle cost of such a policy.

Lemma 4.3. *Let (S, τ) be the policy that solves the optimal stopping problem given a reward g. If the expected cycle cost of this policy is negative (positive), then its average cycle cost*

$$\frac{K + \mathbb{E}_S\left[\sum_{n=1}^{\tau} L(I_n)\right]}{\mathbb{E}_S[\tau]}$$

is smaller (larger) than g.

Proof. Suppose that the average cycle cost $v(S)$ is negative. Then we have from (2) that

$$\begin{aligned} 0 > v(S) &= K + \mathbb{E}_S\left[\sum_{n=1}^{\tau}(L(I_n) - g)\right] \\ &= K + \mathbb{E}_S\left[\sum_{n=1}^{\tau} L(I_n)\right] - g\mathbb{E}_S[\tau] \end{aligned}$$

which implies that the long-run average cost is smaller than g. The same reasoning applies to the opposite claim. □

Lemma 4.3 immediately translates into a method for finding a policy with minimum average cycle cost. It is evident that the expected cycle cost is decreasing in g. Hence, by solving a series of optimal stopping problems with different rewards one can converge to a reward g^* for which the expected cycle cost is zero. The average cycle cost of this policy will be exactly g^*. We discuss this procedure further at the end of this section.

The policy that minimizes the average cycle cost can be used cycle after cycle. It is clear that the long-run average cost of the policy is independent of the starting inventory level, as the inventory process reaches the stopping set from any inventory level in finite time at finite cost. Then, we have from the renewal-reward theorem that its long-run average cost is equal to its average cycle cost. We do not yet know whether this policy is the long-run average cost optimal policy that solves (1). Breiman (1964, Section 10.14) shows that this is indeed the case.

Lemma 4.4. *Let g^* be the reward such that the expected cycle cost of the policy that solves the associated optimal stopping problem is zero. Then, this policy is also the long-run average cost optimal and its long-run average cost is g^*.*

Proof. In Breiman's (1964, Section 10.14) terminology, the inventory control problem (1) is a binary decision renewal problem with origin S. The cost to return to the origin is K, the incentive fee is $-L(i)$, and the time of return to the origin is 0. With this, the claim directly follows from Breiman (1964, Theorem 10.5). □

We can now establish the optimality of (s, S) policies as an immediate consequence of the result presented above.

Theorem 4.5. *The long-run average cost optimal policy is an (s, S) policy.*

Proof. It is sufficient to show that the policy that solves the optimal stopping problem for any given reward g is an (s, S) policy. The optimal policy always initiates a cycle at the same inventory level in S in C, and it stops whenever the inventory process hits D. This is clearly an (s, S) policy with s as the right boundary of D, and S as the minimizer of v on C. □

We have thus far achieved the first two objectives mentioned in the previous section. We now consider the third objective and devise an efficient procedure to compute the optimal policy parameters. The procedure proceeds with the steps described below.

1. Define bounds g_- and g_+ such that the optimal long-run average cost g^* is an element of $(g_-, g_+]$. For instance, $g_- = \min L(i)$ and $g_+ = g_- + K$ are sensible options because if $g < g_-$ the reward is smaller than the minimum achievable cost per period and if $g > g_+$ the reward is larger than the average cost under the optimal base stock policy.
2. Use a root finding algorithm to choose $g \in (g_-, g_+]$.
3. For the current g, identify D, C, and, s as in (3) and (4). Compute $v(\cdot)$ on C with the dynamic programming Eq. (6) and find S where it attains the minimum.
4. Terminate if $v(S)$ is sufficiently close to zero (for a more refined termination criterion see Corollary 4.7). Otherwise, if $v(S) < 0$ set $g_+ = g$ and if $v(S) > 0$ set $g_- = g$, and return to Step 2.

The procedure outlined above is computationally efficient as it converges exponentially fast and requires only a line search over a limited domain in each iteration. It is also conceptually simple and can be coded in a few lines.

Finally, we establish a result that leads to a simple and more efficient termination criterion for our procedure.

Lemma 4.6. *Suppose there exists an optimal policy that solves the optimal stopping problem for rewards g_- and g_+ with $g_- < g_+$. Then this policy also solves the optimal stopping problem for any $g \in (g_-, g_+)$.*

Proof. Because the optimal policies are the same for the rewards g_- and g_+, the continuation set C must be the same for g_- and g_+. Then, it follows from (3) that C is also the same for any $g \in (g_-, g_+)$. This in turn implies that the stopping time τ is the same for $g \in [g_-, g_+]$. Hence, it suffices to show that the value function is minimized at the same S for all rewards on $[g_-, g_+]$.

To denote the dependence on the reward $g \in (g_-, g_+)$, we use a subscript and write the value function as $v_g(i) = K + \mathbb{E}_i\left[\sum_{n=1}^{\tau}(L(I_n) - g)\right]$, where τ is the (common) optimal stopping time. Then, we see that

$$\begin{aligned} v_{g_-}(i) - v_g(i) &= (g - g_-)\mathbb{E}_i[\tau], \\ v_{g_-}(i-1) - v_g(i-1) &= (g - g_-)\mathbb{E}_{i-1}[\tau]. \end{aligned}$$

If we subtract these equations, we obtain

$$v_{g_-}(i) - v_{g_-}(i-1) = v_g(i) - v_g(i-1) + (g - g_-)(\mathbb{E}_i[\tau] - \mathbb{E}_{i-1}[\tau]).$$

Observe that $g - g_- > 0$ and $\mathbb{E}_i[\tau] - \mathbb{E}_{i-1}[\tau] \geq 0$. Therefore, we have $v_{g_-}(i) - v_{g_-}(i-1) \geq v_g(i) - v_g(i-1)$ for all $i \in C$. Also, by assumption, S is a minimizer of v_{g-}, hence, $v_{g-}(S-1) \geq v_{g-}(S)$, which in turn implies that $0 \geq v_{g-}(S) - v_{g-}(S-1)$. Combining this with the previous inequality, we obtain

$$0 \geq v_{g_-}(S) - v_{g_-}(S-1) \geq v_g(S) - v_g(S-1),$$

which immediately shows that $S-1$ cannot be a minimizer of v_g.

We can use similar reasoning to see that as S is a minimizer for v_{g_+}, then $S+1$ cannot be the minimizer for v_g. □

Corollary 4.7. *If the lower and upper bounds of g yield the same policy, then this policy is optimal.*

5. Discussion

In this paper, we showed that simple methods to establish the existence and characterization of optimal policies and efficient numerical procedures to compute optimal policy parameters could have been developed much earlier if stochastic inventory problems were looked through the lens of optimal stopping theory, based on Breiman's (1964) results. To that end, we focused on the classical periodic-review stochastic inventory problem with convex inventory costs under the long-run average cost criterion—a problem which has been subjected to detailed scrutiny (see e.g. Beyer & Sethi, 1999; Feng & Xiao, 2000; Iglehart, 1963; Tijms, 1986; Veinott & Wagner, 1965; Zheng, 1991; Zheng & Federgruen, 1991). We now conclude with a brief discussion on how the presented approach relates to the existing literature.

Our approach combines and builds upon concepts of renewal-reward processes, optimal stopping, dynamic programming, and root-finding; yet it follows a rather simple recipe. That is, decompose the overall inventory problem into replenishment cycles—which is possible due to the renewal-reward structure of the problem—and formulate the sub-problem associated with a single replenishment cycle as a parametric optimal stopping problem. The cycle decomposition and parametric optimization ideas have already been used for the very problem we consider in this study. Zheng (1991) uses cycle decomposition in his proof of the optimality of (s, S) policies. Feng & Xiao (2000) employ cycle decomposition and parametric optimization in their method of finding optimal (s, S) policies. Their computational method is perfectly in line with the procedure outlined in the current manuscript, despite they do not depart from the optimal stopping characterization of the underlying problem. They also present further algorithmic refinements and conduct a detailed comparative study that shows their method outperforms Zheng & Federgruen's (1991) well-known algorithm with respect to computational efficiency. Like the aforementioned studies, we rely on the renewal-reward structure of the inventory system. The point we deviate is formulating the parametric problem as an optimal stopping problem. This enables us to directly use Breiman's (1964) results in establishing the optimality of (s, S) policies and computing the optimal policy parameters. Bell (1970) considers the discounted cost counterpart of the problem and uses a very similar approach to ours where the sub-problem is formulated as an optimal stopping problem. This permits him to limit the search space of optimal policy parameters and improve on Veinott & Wagner's (1965) algorithm. In contrast, we iteratively solve a sequence of optimal stopping problems which eventually lead to the optimal policy. There are also studies that use cycle decomposition and parametric optimization to establish proofs and computational methods of optimal policies for problems beyond the periodic-review stochastic inventory problem with convex inventory costs. For instance, Chen & Feng (2006) consider inventory problems with non-quasi-convex costs, Feng & Chen (2011) focus on joint inventory and pricing problems, Foreest & Wijngaard (2014) address production-inventory systems. Germs et al. (2016) consider production-clearing systems. In all these studies, the analysis rests on exploiting the renewal-reward structure of the underlying inventory system, and follows by cycle decomposition and parametric optimization where the sub-problem is tailored to account for the specifications of the problem.

We note that there also are other studies that approach inventory problems with methods of optimal stopping (see e.g. Berling & Martínez-de Albéniz, 2011; Frenk et al., 2019; Oh & Özer, 2016; Ozyoruk et al., 2022; Shi & Liu, 2020; Weiss, 1980). We do not provide a detailed account of these since they have little resemblance to our work as they do not use optimal stopping in conjunction with the renewal-reward structure of the inventory systems.

References

Bell, C. E. (1970). Improved algorithms for inventory and replacement-stocking problems. *SIAM Journal on Applied Mathematics, 18*(3), 558–566.

Berling, P., & Martínez-de Albéniz, V. (2011). Optimal inventory policies when purchase price and demand are stochastic. *Operations Research, 59*(1), 109–124.

Beyer, D., Cheng, F., Sethi, S. P., & Taksar, M. (2010). *Markovian demand inventory models*. Springer.

Beyer, D., & Sethi, S. P. (1999). The classical average-cost inventory models of Iglehart and Veinott–Wagner revisited. *Journal of Optimization Theory and Applications, 101*(3), 523–555.

Breiman, L. (1964). Stopping-rule problems. In E. Beckenbach (Ed.), *Applied combinatorial mathematics* (pp. 284–319). John Wiley and Sons. chapter 10

Chen, F. Y., & 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), 287–306.

Feng, Y., & Chen, Y. (2011). A computational approach for optimal joint inventory-pricing control in an infinite-horizon periodic-review system. *Operations Research, 59*(5), 1297–1303.

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.

Foreest, N. D. V., & 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.

Frenk, J. B. G., Javadi, S., & Sezer, S. O. (2019). An optimal stopping approach for the end-of-life inventory problem. *Mathematical Methods of Operations Research, 90*, 329–363.

Germs, R., Foreest, N. D. V., & Kilic, O. A. (2016). Optimal control policies for production-clearing systems under continuous review. *European Journal of Operational Research*, 747–757.

Iglehart, D. L. (1963). Dynamic programming and stationary analyses of inventory problems. In H. Scarf, D. Gilford, & M. Shelly (Eds.), *Multistage inventory models and techniques* (pp. 1–31). Stanford, California: Stanford University Press.

Oh, S., & Özer, O. (2016). Characterizing the structure of optimal stopping policies. *Production and Operations Management, 25*(11), 1820–1838.

Ozyoruk, E., Erkip, N. K., & Ararat, Ç. (2022). End-of-life inventory management problem: Results and insights. *International Journal of Production Economics, 243*, 108313.

Scarf, H. E. (1959). The optimality of (S, s) policies in the dynamic inventory problem. In K. Arrow, S. Karlin, & H. E. Scarf (Eds.), *Mathematical methods in the social sciences* (pp. 196–202). Stanford: Stanford University Press. chapter 13

Shi, Z., & Liu, S. (2020). Optimal inventory control and design refresh selection in managing part obsolescence. *European Journal of Operational Research, 287*(1), 133–144.

Tijms, H. C. (1986). *Stochastic modelling and analysis, a computational approach*. J. Wiley & Sons.

Veinott, A. F., & Wagner, H. M. (1965). Computing optimal (s, S) inventory policies. *Management Science, 11*(5), 525–552.

Weiss, H. J. (1980). Optimal ordering policies for continuous review perishable inventory models. *Operations Research, 28*(2), 365–374.

Zheng, Y. S. (1991). A simple proof of the optimality of (s, S) policies in infinite-horizon inventory systems. *Journal of Applied Probability, 28*, 802–810.

Zheng, Y. S., & Federgruen, A. (1991). Finding optimal (s, S) policies is about as simple as evaluating a single policy. *Operations Research, 39*(4), 654–665.