Inventory management problems with periodic and controllable resets occur in the context of managing water storage in the developing world and retailing limited-time availability products. In this paper, we consider a set of sequential decision problems in which the decision-maker must not only balance holding and shortage costs but discard all inventory before a fixed number of decision epochs, with the option for an early inventory reset. Finding optimal policies using dynamic programming for these problems is particularly challenging since the resulting value functions are non-convex. Moreover, this structure cannot be easily analyzed using existing extended definitions, such as $K$-convexity. Our key contribution is to present sufficient conditions that ensure the optimal policy has an easily interpretable structure that generalizes the well-known $(s,S)$ policy from the operations literature. Furthermore, we demonstrate that the optimal policy has a four-threshold structure under these rather mild conditions. We then conclude with computational experiments, thereby illustrating the policy structures that can be extracted in several inventory management scenarios.

Key words: supply chain management, inventory theory and control, dynamic programming, health care management, humanitarian operations

1. Introduction

The study of inventory management with stochastic demand has been a key discipline in operations management since the inception of the field. In general, these problems consider a decision-maker who must choose either an order quantity or a replenishment level such that it minimizes their
overall expected supply chain costs or that they are able to maintain a certain level of service with high probability. Several models have been proposed in the literature to address this setting, such as the newsvendor model for single period replenishment (Arrow et al. 1951), the \((r, Q)\) model for continuous review policies (Hadley and Whitin 1963), and the \((s, S)\) model for periodic inventory review models (Scarf 1959). These settings generally focus on finding an optimal inventory policy, that is, a decision rule that maps the current inventory level of the facility to an order quantity. To find this policy, classical models have mainly focused on balancing holding costs (the costs incurred from holding excess inventory) and shortage costs (the costs associated with not having enough units on hand to satisfy demand). However, when products are perishable or subject to strict health and safety regulations, the assumptions that underlie these policies are no longer met.

Many settings, such as food procurement (van Donselaar et al. 2006, Blackburn and Scudder 2009, Farahani et al. 2012), medical supply chain management (Prastacos 1984, Pierskalla 2005, Shen et al. 2011), and certain seasonally sensitive retail supply chains (Caro and Gallien 2007, Nagurney et al. 2013), require managing perishable inventory and therefore do not satisfy the classical settings of the aforementioned inventory management models. The common way this constraint is addressed is through the use of strategic discarding policies. Essentially, the decision-maker is able to form a ranking of which products are most valuable to keep for longer durations and to individually discard expired units based on costs and health constraints. While these policies are applicable to many different contexts, if the particular goods being managed are commingled or not easily separable, it is not feasible to strategically discard individual units, but instead, the entirety of the inventory must be discarded. For example, in the case of residential water storage, as new water is purchased, it mixes with older water that is more likely to have been tainted and cannot be readily separated from the previous water in the tank.

In this paper, we propose new models and techniques to address the setting of a perishable inventory with known expiration time, where strategic discarding is not possible. In particular, we propose to model these systems as inventory control problems with strict reset constraints. In
this setting, at each epoch, the decision-maker is faced with two decisions: first, whether or not to discard the entire inventory on hand; and second, whether or not to order inventory up to a certain level. Since the inventory is set to expire at a predetermined time, the decision-maker must discard the entirety of the inventory after a fixed number of epochs have passed. The main challenge in this problem is to find an inventory policy for the decision-maker that balances safety constraints, in conjunction with holding and shortage costs, by making refresh and reorder decisions.

1.1. Applications

The setting of mixing perishable inventory with reset decisions is common in many real-world applications. Here, we present two settings that exemplify the assumptions of the model. First, we consider a non-profit example of managing water storage in a residential building in the developing world. Next, we describe a profit example in the case of a retail supply chain with switching product lines.

1.1.1. Non-Profit: Water Storage Problem

Over 300 million people around the world have an intermittent water supply (Kumpel and Nelson 2016), making water availability a pervasive public health crisis. This problem is more severe in the developing world, where the current distribution infrastructure is not capable of providing cities with continuous water supplies. Consequently, most households resort to communal and personal water storage containers to maintain a water supply throughout the day, and these water storage containers are filled once every few days during intermittent periods of water availability (World Bank 2010, WHO and UNICEF 2017). However, water stored in these containers is prone to contamination, and water stored long-term is likely to have higher concentrations of bacterial and viral pathogens (Gadgil 1998, Zérah 1998, Coelho et al. 2003, Tokajian and Hashwa 2003, Lee and Schwab 2005, Klingel 2012). The problem of managing local water storage can be posed in the framework of perishable inventory that we have described previously. Each day, the decision-maker must choose how much additional water to purchase to satisfy the demand for that day. Since the longer water sits in the tank, the higher its chance of getting contaminated, this can be modeled as a holding penalty to ensure that water
does not get stored for too long. Likewise, if the amount of water purchased is lower than the
demand realized for that day, expediting the shipment of additional water may be costly, which can
be modeled as a shortage penalty. Finally, since the risk of contamination makes water perishable
and contaminated water cannot be easily separated from clean water, the only option available to
the decision-maker is a “reset” type action where the whole amount of water on hand is discarded.

1.1.2. Profit: Retail Management Problem In many retail supply chain settings, product
demand is highly dependent on seasonal tastes or fashion trends. This can affect various types of
products, ranging from apparel and food to consumer electronics [Matsuo and Ogawa 2007, Caro
and Martínez-de-Albéniz 2015]. In the specific case of clothing, this has previously been studied
in the operations literature through the lens of fast fashion [Barnes and Lea-Greenwood 2006,
Caro and Gallien 2010, Bhardwaj and Fairhurst 2010, Cachon and Swinney 2011]. The key to
managing these inventories is that as new trends arrive, new product lines must be introduced
to meet the most current customer needs. While production and inventory costs may remain the
same for new product lines, as products fall out of fashion with consumers, the old product line
must be completely discarded to make room for the new product line. This setting can again
be modeled by our perishable inventory with reset control framework. Although, unlike water,
apparel and similar products are easy to separate, due to how tastes shift in terms of seasonality
and fashion, all products become obsolete simultaneously and not individually. This means that
strategic discarding is not possible since the entirety of the stock needs to be replaced. Hence, the
key challenge to designing an inventory policy in this setting is similar to that of water storage. Not
only do inventory ordering, shortage, and holding costs need to be balanced, but also the decision-
maker must consider at what point it is advantageous to take a “reset” action corresponding to
introducing a new product line.

1.2. Related Literature

In this paper, we examine the problem of perishable inventory with reset control, a special case of
multi-period inventory management with periodic review and stochastic demand. In the context
of supply chain theory, a policy is a function that maps the current state of our inventory into an ordering decision (Snyder and Shen 2019). For periodic-review inventory models with stochastic demands, a base-stock policy (e.g., newsvendor model) or an \((s, S)\) policy is an example of inventory control policy. The fundamental idea underlying the \((s, S)\) policy is the following: in each time period, we observe the current inventory position—if the inventory position falls below \(s\), then we place an order of sufficient size to bring the inventory position to \(S\), where the quantity \(s\) is known as the reorder point and \(S\) as the order-up-to level (Arrow et al. 1951, Scarf 1959). To this classical framework, we seek to add a reset control action to empty the entire inventory and generalize the \((s, S)\) policy for such problems. Mintz et al. (2017) studied this approach numerically, and our contributions include the theoretical analysis of the optimal policy in the presence of periodic and controllable resets. In addition to these works, we draw on two main streams of literature for the basis of our modeling and analysis, namely inventory management and dynamic programming.

1.2.1. Inventory and Supply Chain Management. The problem we describe in this paper is closely related to the setting of managing perishable inventory (Pierskalla and Roach 1972, Nahmias 1982, Hsu 2000, Hsu and Lowe 2001, Chu et al. 2005, Karaesmen et al. 2011, Coelho and Laporte 2014), which has been explored primarily in the context of healthcare settings. In these perishable inventory models, other than the cost of production, costs arise either from not having enough inventory to meet demand (shortage costs) or from holding excess inventory (holding costs) that must be discarded due to the perishable nature of the good. In contrast to our setting, this stream of literature generally assumes that expired units of inventory can be disposed of individually, which allows salvage and disposal costs to be incorporated into the holding cost, and hence, policies of strategic inventory removal can be employed (Rosenfield 1989, 1992). In our setting, however, decision-makers cannot dispose of single units of their stock strategically. For example, in the water storage problem, fresh water may mix with still water, and they cannot be easily separated. Instead, the decision-maker must decide to either keep or dispose of their entire inventory in order to reduce the risk of drinking contaminated water. Similarly, in the retail
management problem, we determine if an entire product line should be kept or discontinued. In the case where the product line is discontinued, all of its inventory must be discarded to free up space for the inventory of the new product line.

1.2.2. Dynamic Programming and Optimal Control It is well known from the operations literature that periodic-review inventory management problems can be modeled as continuous-state dynamic programs (Scarf 1959, Bertsekas 1995). For most of these models, while it is possible to derive structural results for the functional form of the optimal policy, it is often difficult to find a closed-form solution for the parameters of the policy. This means that, in general, numerical algorithms must be used to find the relevant parameter values. The main two numerical methods developed to solve these dynamic programming problems are value and policy iteration (Bertsekas 1995), which require finding a fixed point in either the value function or the policy space, respectively. Essentially, this means algorithms are attempting to find fixed points in functional spaces that may be infinite dimensional. Even in finite high-dimensional state spaces, the exact calculation for these approaches is numerically difficult, an effect commonly referred to as the curse of dimensionality. To address this challenge, a number of approximate dynamic programming approaches (Bertsekas 1995, de Farias and Van Roy 2003, Powell 2007, Ryzhov et al. 2012, Kariotoglou et al. 2013, Haskell et al. 2016) have been developed. The key to these approaches is to perform a value or policy iteration with an inexact, but tractable, representation of the value function or policy. However, the convergence rate of (exact or approximate) value and policy iteration is governed by the discount factor, and so convergence is slow for discount factors close to 1 (Bertsekas 1995, Zidek and Kolmanovsky 2016). Due to the particular structure of the problem we examine in our setting, we propose the use of the Binary Dynamic Search algorithm (BiDS), originally devised by Mintz et al. (2017). In contrast to value and policy iteration that operate in function spaces, the BiDS algorithm finds a fixed point in a vector space using binary search. Since the state space is small for water storage and retail management problems, our numerical results solve the exact dynamic program; however, in principle, the BiDS algorithm could also be used for approximate dynamic programming by using inexact representations of the value function.
1.3. Outline and Contributions

Our main contribution in this paper is to study the structure of inventory management problems with periodic review and reset control. As part of this analysis, we theoretically characterize and prove the structure of the optimal policy for these inventory management problems. Our analysis shows that the resulting policy is a threshold policy that can be seen as an extension of the classic \((s,S)\) policy. In addition, our analysis provides new theoretical methods for analyzing inventory policies since the non-convexity induced by our problem structure is distinct from that present in classical inventory models.

In addition to studying the structural properties of the problem, our second contribution in this paper is to employ a novel algorithm for solving stochastic optimal control problems with the specific structure of controlled resets to a single state and with constraints on the maximum time in between system resets. Unlike value and policy iteration, which require finding a fixed point in an infinite-dimensional functional space \cite{Bertsekas1995}, we implement a Binary Dynamic Search (BiDS) algorithm \cite{Mintz2017} that converts the control problem into finding a fixed point in a vector space using binary search. Although value and policy iteration converge exponentially, the convergence rate can be practically slow when the discount factor is close to 1 \cite{Bertsekas1995,Zidek2016}. The BiDS algorithm uses binary search and can thus compute the value functions of problems with controlled resets using substantially less computations. We apply BiDS to numerically solve this particular problem and experimentally verify our structural results. In doing so, we generalize the previous \((s,S)\) inventory policy to a new threshold structure that incorporates reset control. To explore the broad applicability of our new structure, we demonstrate that we have designed a policy that can be solved, interpreted, and implemented in both profit and non-profit operations.

The rest of the paper proceeds as follows. We first describe the stochastic optimization problem for the general reset control model in Section 2, with applications to water storage problem in Section 2.1 and retail management problem in Section 2.2. We then discuss the structural properties
of the optimal policy and the sufficient conditions that ensure the threshold structure in Section 3, along with the numerical algorithm and its results from both water and retail problems in Section 4.

2. Reset Control Formulation and Applications

In this section, we introduce a broad class of stochastic optimal control problems with controlled resets to a single state and with constraints on the maximum time span in between resets of the system. Let the subscript \( n \in \mathbb{Z}_+ \) denote the index of the decision epoch, and consider the discrete-time dynamical system

\[
\begin{align*}
x_{n+1} &= h(x_n, \tau_n, u_n, w_n) \\
t_{n+1} &= \tau_n + 1 \\
\xi_n &= x_n \cdot (1 - r_n) + \zeta \cdot r_n \\
\tau_n &= t_n \cdot (1 - r_n)
\end{align*}
\]

where \( x_n \times t_n \in \mathbb{R}^{n_x} \times \mathbb{Z}_+ \) are states, \( \xi_n \times \tau_n \in \mathbb{R}^{n_x} \times \mathbb{Z}_+ \) are pseudo-states, \( u_n \times r_n \in \mathbb{R}^{n_u} \times \mathbb{B} \) are control actions, disturbance terms \( w_n \in \mathbb{R}^{n_w} \) are i.i.d. random variables, and \( h : \mathbb{R}^{n_x} \times \mathbb{Z}_+ \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \rightarrow \mathbb{R} \) is a deterministic function. The interpretation is that the control \( r_n = 1 \) resets the system to a known initial state \( \zeta \in \mathbb{R}^{n_x} \), the state \( t_n \) keeps track of how many time steps have passed since the last system reset, and the function \( h \) describes the dynamics when there is no reset. The inclusion of pseudo-states \( \xi_n \) and \( \tau_n \) captures the reset dynamics since these terms are set to the known initial states when \( r_n = 1 \).

Given the discount factor \( \gamma \in [0, 1) \), our goal is to solve the following stochastic control problem:

\[
\begin{align*}
\min \mathbb{E} \left[ \sum_{n=0}^{\infty} \gamma^n \left( g(\xi_n, \tau_n, u_n, w_n) + s(x_n, t_n, w_n) \cdot r_n \right) \right], \\
\text{s.t. (1), } t_n \leq k, u_n \in \mathcal{U}(\xi_n, \tau_n, w_n), \quad \text{for } n \geq 0.
\end{align*}
\]

where \( g : \mathbb{R}^{n_x} \times \mathbb{Z}_+ \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}_+ \) is a non-negative and continuous stage cost, and \( s : \mathbb{R}^{n_x} \times \mathbb{Z}_+ \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}_+ \) is a non-negative and continuous reset cost with \( s(\zeta, t, w) \equiv 0 \). The constraint \( u_n \in \mathcal{U}(\xi_n, \tau_n, w_n) \) restricts the possible control actions to lie in a set \( \mathcal{U}(\xi_n, \tau_n, w_n) \), and the constraint
\( t_n \leq k \) requires the system be reset at least once every \( k \) time steps. For notational convenience, we will simply refer to the set \( U_n := U(\xi_n, \tau_n, w_n) \).

While this stochastic control formulation is quite general, we now present two instantiations of this model in real-world settings. First, we discuss a non-profit example of water storage control, and then we describe a for-profit example of retail management with changing product lines.

### 2.1. Example: Water Storage Problem

In this section, we consider the stochastic inventory control problem of water management in the developing world, where a single decision-maker must maintain the level of potable water in a residential water tank. Unlike water management in the developed world, we assume the residence does not have access to a continuous source of water, and thus it must be purchased in bulk at the beginning of the day from a communal source. Every few days, the tank needs to be fully emptied so that it can be cleaned to eliminate pathogen growth. Therefore, in this setting, the decision-maker has two actions they can take each day: first, decide if they would like to purge the tank, and second, decide how much water to purchase at the beginning of the day. These decisions must be made in such a way that optimally trades off the financial cost of purchasing water or expediting purchases when the amount is short, and the implicit health costs of letting water sit longer in the tank (that can be thought of as a holding cost).

This problem can be modeled as a stochastic control problem that is a special case of the formulation presented in Section 2. For day \( n \), let \( x_n \in \mathbb{R}_+ \) be the state variable representing the amount of water stored in the tank, and let \( t_n \in \mathbb{Z}_+ \) be the state variable representing the number of days since the tank was last emptied. Let the decision-maker’s actions of the amount of water to purchase at the start of the day and whether or not to empty the tank be represented by \( u_n \in \mathbb{R}_+ \) and \( r_n \in \mathbb{B} \), respectively. We model the demand for each day as the i.i.d. random disturbance process \( w_n \in \mathbb{R}_+ \).

Let the dynamics of the system be described by the following:

\[
\begin{align*}
x_{n+1} &= (x_n \cdot (1 - r_n) + u_n - w_n)^+, \\
t_{n+1} &= t_n \cdot (1 - r_n) + 1,
\end{align*}
\]
where \((x)^+ = \max\{x, 0\}\) and \((x)^- = \min\{x, 0\}\). Here, \((3)\) states that the level of water at day \(n + 1\) will equal the current level of water at day \(n\) plus any amount of water purchased that day, minus the amount of water demanded, but cannot go below zero. If the tank is flushed on day \(n\), then \((3)\) states that the water level of the tank at day \(n + 1\) will not depend on the amount of water at day \(n\). Likewise, \((4)\) states that the number of days since the tank has been flushed increments by 1 each day until the day it is flushed, in which case the count resets to 1. Let \(q : \mathbb{R}_+ \to \mathbb{R}_+\) be a non-negative and non-decreasing function that represents the increased health costs associated with letting water sit in the tank for additional days. Let \(c_u, c_r, p \in \mathbb{R}_+\) represent the variable cost of purchasing water, the per-unit cost of purging the tank, and the shortage penalty for not having sufficient water to meet demand. Using these cost parameters, state variables, dynamics equations, and discount factor \(\gamma \in [0, 1)\), the decision-maker’s problem can be formulated as follows:

\[
\min \mathbb{E}\left[ \sum_{n=0}^{\infty} \gamma^n \left( c_u u_n - p \cdot (\xi_n + u_n - w_n)^- + q(\tau_n) \cdot (\xi_n + u_n - (\xi_n + u_n - w_n)^+) + c_r x_n r_n \right) \right],
\]

\[\text{s.t. } x_{n+1} = (\xi_n + u_n - w_n)^+, \quad \text{for } n \geq 0,\]

\[
t_{n+1} = \tau_n + 1, \quad \text{for } n \geq 0,
\]

\[
\xi_n = x_n \cdot (1 - r_n), \quad \text{for } n \geq 0,
\]

\[
\tau_n = t_n \cdot (1 - r_n), \quad \text{for } n \geq 0,
\]

\[
t_n \leq k, \quad \text{for } n \geq 0,
\]

\[
r_n \in \mathbb{B}, \quad \text{for } n \geq 0,
\]

\[
u_n \in [0, c_{max} - \xi_n], \quad \text{for } n \geq 0.
\]

For this formulation, the state space is augmented to include pseudo-states \(\xi_n, \tau_n\) that represent intermediate values of water in the tank and time since the last reset, given the reset action on day \(n\), respectively. Using these states, the objective terms \(p \cdot (\xi_n + u_n - w_n)^-\) and \(q(\tau_n) \cdot (\xi_n + u_n - (\xi_n + u_n - w_n)^+)\) represent the total shortage cost and health costs at day \(n\), respectively, the interpretation being that the decision-maker pays shortage costs only when demand exceeds water supply and that water quality deteriorates only if a surplus remains in the tank. The quantity
\((\xi_n + u_n - (\xi_n + u_n - w_n)^+)\) is the amount of water consumed on the \(n\)-th day because \(\xi_n + u_n\) is the amount of water available at the beginning, and \(\xi_n + u_n - w_n\) is the amount of water that is still unused at the end of the day. Since the function \(q\) is assumed to be monotonically increasing, the \(q(\tau_n)\) term indicates that the quality of water deteriorates as time passes between emptying the tank and also ensures that consuming water that has been stored for longer durations of time is more heavily penalized. Likewise, the terms \(c_u u_n\) and \(c_r x_r n\) represent the total costs of purchasing water and flushing the tank, respectively, where the flushing cost is only incurred if the reset action is taken. The constraint (10) ensures that the tank is purged at least once every \(k\) days, and the constraint (12) ensures that water is not purchased in excess of the tank capacity \(c_{max}\).

A defining feature of this problem is that instead of considering the holding/shortage cost trade-off, we formulate the problem based on how long water has been stored and how much water has actually been consumed. In circumstances where people do not have a continuous supply of water and thus have no other choice but to store water in a local storage container that lacks disinfection capabilities, we are taking an optimization approach to decide when to drain the water tank and how much water to fill it when available, with the objective of reducing the risk of contamination and shortage. In Section 4, we will solve this problem numerically to design an efficient, interpretable policy for managing water storage systems so that it is easily implementable via a reference table that can be widely distributed to the public through paper pamphlets or the internet.

### 2.2. Example: Retail Management Problem

Next, we consider the setting of retail inventory management in the case of changing product lines. In this setting, the decision-maker is tasked with managing the inventory of limited-time availability product lines and must decide the quantity of stock to store and the timing for changing to a new product line such that they are able to minimize inventory costs. This is accomplished by balancing the trade-offs between the ordering cost, the holding cost for storing the inventory, the shortage cost, the stock wastage from emptying the inventory, and the cost of switching the product line.
Like the water storage problem, this retail inventory problem can also be modeled as a special case of the reset control problem presented in Section 2. For each week \( n \), let the state variables \( x_n \in \mathbb{R}_+, t_n \in \mathbb{Z}_+ \) represent the amount of inventory in stock of the current product line at the beginning of the week, and the number of weeks since the current product line has been offered. Let the decision-maker’s weekly control actions of how much new inventory to add and whether or not to replace the current product line with a new one be given by \( u_n \in \mathbb{R}_+, r_n \in \mathbb{B} \), respectively. Much like the water storage case, we model demand as an i.i.d. disturbance process \( w_n \in \mathbb{R}_+ \). The state dynamics for this model are identical to those in (3) and (4), except that the interpretation of (3) now reflects that the inventory level increases with each additional product purchased and decreases by the realization of weekly demand (or the taking of a reset action) but cannot go below zero, while (4) reflects the age of the current product line. Let \( c_u, k_u > 0 \) be the variable and fixed ordering costs associated with the current product line, respectively, and \( c_r, k_r > 0 \) be the associated variable and fixed ordering costs of purchasing units from the new product line. Furthermore, let \( p, q > 0 \) be the per-unit shortage and holding costs, respectively. Then using these costs, states, dynamics, and discount factor \( \gamma \in [0,1) \), the decision-maker’s problem can be formulated as follows:

\[
\min \mathbb{E}\left[ \sum_{n=0}^{\infty} \gamma^n \left( (c_u u_n + k_u)1_{\mathbb{R}_+^+}(u_n) - p(\xi_n + u_n - w_n) - q(\xi_n + u_n - w_n)^+ + (c_r x_n + k_r) r_n \right) \right] \tag{13}
\]

s.t. \( x_{n+1} = (\xi_n + u_n - w_n)^+ \), for \( n \geq 0 \)
\[
t_{n+1} = \tau_n + 1, \quad \text{for} \quad n \geq 0 \tag{14}
\]
\[
\xi_n = x_n \cdot (1 - r_n), \quad \text{for} \quad n \geq 0 \tag{15}
\]
\[
\tau_n = t_n \cdot (1 - r_n), \quad \text{for} \quad n \geq 0 \tag{16}
\]
\[
t_n \leq k, \quad \text{for} \quad n \geq 0 \tag{17}
\]
\[
r_n \in \mathbb{B}, \quad \text{for} \quad n \geq 0 \tag{18}
\]
\[
u_n \in [0, c_{\max} - \xi_n], \quad \text{for} \quad n \geq 0 \tag{19}
\]
where the function \( 1_{\mathbb{R}^+} : \mathbb{R} \mapsto \mathbb{B} \) is defined as the indicator
\[
1_{\mathbb{R}^+}(u) = \begin{cases} 
1, & \text{if } u > 0, \\
0, & \text{otherwise.}
\end{cases}
\] (21)

The total cost incurred in the \( n \)-th week comprises up to four components: (i) the purchasing cost \( c_u u_n + k_u \), where \( c_u \) is the cost per unit ordered and \( k_u \) is the fixed cost associated with a positive inventory order, (ii) the shortage cost \( p \cdot (\xi_n + u_n - w_n)^- \) that represents the loss incurred when demand is unmet, (iii) the holding cost \( q \cdot (\xi_n + u_n - w_n)^+ \) for having too much inventory relative to the actual demand, and (iv) the reset cost \( c_r x_n + k_r \), where \( c_r \) is the waste penalty per unit discarded and \( k_r \) is the fixed cost associated with resetting the product line. The constraint (18) ensures that the product line is fully changed at least once every \( k \) weeks, and the constraint (20) ensures that the inventory capacity \( c_{\text{max}} \) is not exceeded.

What differentiates this inventory management problem from the classic \((s,S)\) setting is the inclusion of reset controls. When the reset control action is taken at time \( n \) (i.e., \( r_n = 1 \)), all products are removed, and the system is reverted to the state \((x_n, t_n) = (0,0)\). This means that by solving this problem, the firm can change product lines earlier than planned if it is advantageous to do so. In other words, given cost parameters, this problem solves the best business strategy tailored for limited-time products, which encompasses not only the optimal inventory (i.e., how much stock to order each period) but also the optimal timing (i.e., when to switch the product line).

3. Structural Results

In this section, we analyze the structure of the reset control problem (2). We begin by describing the dynamic programming equations first. Let \( J : \mathbb{R}^n \times \mathbb{Z}_+ \mapsto \mathbb{R}_+ \) be the optimal cost-to-go function, that is, \( J(x,t) \) is defined as the minimum value of (2) for the initial conditions \( x_0 = x \) and \( t_0 = t \). Let \( J_0 = J(\zeta,0) \), the cost-to-go from the reset state. Then the dynamic programming equations can be characterized using the following result from Mintz et al. (2017).
Proposition 1. (Mintz et al. 2017) The dynamic programming equations for (3) are given by

\[
J(\zeta, 0) = \min_{u \in U_0} \mathbb{E}\left[ g(\zeta, 0, u, w) + \gamma J(h(\zeta, 0, u, w), 1) \right]
\]

\[
J(x, t) = \min \left\{ J_0 + \mathbb{E}\left( s(x, t, w) \right), \min_{u \in U_t} \mathbb{E}\left[ g(x, t, u, w) + \gamma J(h(x, t, u, w), t + 1) \right] \right\}
\]

\[
J(x, k) = J_0 + \mathbb{E}\left( s(x, k, w) \right)
\]

where the middle \( J(x, t) \) holds for all \( x \) and \( t = 0, \ldots, k - 1 \).

These are the new dynamic programming equations that result from introducing reset control. As shown here, excluding the last period (\( t = k \)) at which the inventory must be reset, we evaluate every period whether resetting is more favorable. The proof for the above equations is found in Mintz et al. (2017).

Our main results in this section will prove that under a set of reasonable assumptions, the optimal fixed point policy \( \pi^* : \mathbb{R} \times \mathbb{Z}_+ \to \mathbb{B} \times \mathbb{R} \) to (22) is a threshold policy. Specifically, we show that it can be characterized by four time-dependant parameters \( s_t, S_t, \sigma_t, \Sigma_t \) and a constant \( \varphi \), and has the following form:

\[
\pi^*(x, t) = \begin{cases} 
  r = 1, u = \varphi, & \text{for } x \in [0, \sigma_t), \\
  r = 0, u = S_t - x, & \text{for } x \in [\sigma_t, s_t), \\
  r = 0, u = 0, & \text{for } x \in [s_t, \Sigma_t), \\
  r = 1, u = \varphi, & \text{for } x \in [\Sigma_t, \infty).
\end{cases}
\]

The intuition behind this policy is that in the region of \([0, \sigma_t)\), there is so little inventory that the cost of resetting is negligible, and so resetting is the optimal decision. In the region of \([\sigma_t, s_t)\), there is enough stock such that the reset cost is too high, and it is likely that the system will experience a shortage, so additional stock must be ordered. In the region \([s_t, \Sigma_t)\), there is enough stock present such that shortages are less likely than excess inventory, so no additional stock is ordered. Finally, the region \([\Sigma_t, \infty)\) indicates that the stock level is so high that it will almost surely all spoil, so it is more beneficial to reset the system.
To prove these results, we will first describe a set of sufficient assumptions that guarantee the structure of the optimal policy. Then we will present the proof for the structure of the policy, and show that it follows from the particular structure of $J$ by induction on the states.

### 3.1. Technical Assumptions

For our structural analysis of the optimal policy, we make the following technical assumptions.

**Assumption 1.** The states are $x_n \in \mathbb{R}_+$, and the state dynamics are $h(\xi_n, \tau_n, u_n, w_n) = (\xi_n + u_n - w_n)^+$. 

**A1** assumes that the inventory is non-negative such that unfilled demand at each stage is not backlogged but rather lost. This translates to the system equation being $x_{n+1} = \max\{0, \xi_n + u_n - w_n\}$, instead of $x_{n+1} = \xi_n + u_n - w_n$. The state dynamics in this assumption represent the remaining inventory from the last period, which is the positive part of the stock quantity stored in the inventory plus the amount of stock added, minus the amount of stock demanded.

**Assumption 2.** The stage cost satisfies

$$
\mathbb{E}(g(\xi_n, \tau_n, u_n, w_n)) = cu_n + K \cdot 1_{\mathbb{R}_+^*}(u_n) + H(\xi_n + u_n, \tau_n)
$$

for a function $H(\cdot, t)$ that satisfies $|\partial_x H(x + u, t)| \leq \kappa_i$ and $\partial_{xx} H(x + u, t) \geq m_i$. Here, $\partial_x H(\cdot, t)$ denotes the first derivative with respect to the first argument, and $\partial_{xx} H(\cdot, t)$ denotes the second derivative with respect to the first argument. The constants $c, K, \kappa_i \geq 0$ are non-negative, and the constants $m_i > 0$ are positive for $t = 0, \ldots, k - 1$.

In **A2**, the stage cost can be decomposed into three parts, where the first term is the ordering cost, the second term is the fixed cost associated with a positive inventory order, and the third term is the function that measures the stage cost (e.g., holding/shortage cost, implying a penalty for both excess inventory and unmet demand at the end of the period). We require the function $H(\cdot, t)$ be strongly convex, and this assumption is key to proving the convexity of the value function with no reset and thereby demonstrating the optimality of the $(s, S)$ policy.
Assumption 3. The known initial state is $\zeta = 0$. The reset cost $R(x,t) := E(s(x,t,w))$ is concave and non-decreasing in $x$. Moreover, its derivative $\partial_x R(\cdot,t)$ is bounded as $|\partial_x R(x,t)| \leq \eta_t$ and is Lipschitz continuous as

$$|\partial_x R(x_n,t_n) - \partial_x R(x'_n,t_n)| \leq L_t \cdot |x_n - x'_n|$$

for non-negative constants $\eta_t, L_t \geq 0$ for $t = 0, \ldots, k - 1$.

A3 assumes that we empty the entire inventory when we reset the system, and that the reset cost is concave and non-decreasing, which reflects the law of diminishing marginal utility. We also require the function $R(\cdot,t)$ to be $\eta_t$-Lipshitz and $L_t$-smooth, which is key to determining the threshold that ensures the optimal policy structure.

Assumption 4. The input constraints are $U_n = \{u : 0 \leq u \leq c_{\max} - \xi_n\}$, where the constant $c_{\max} \geq 0$ is non-negative.

A4 is an input constraint that enforces an upper bound $c_{\max}$ on the level of stock that can be accommodated.

Assumption 5. The $w_n$ are i.i.d., and the density of $w_n$ is given by a function $f(w)$ that satisfies: $f(w) = 0$ for all $w < 0$, is Lipschitz continuous for $w \geq 0$ with a non-negative constant $L$, and bounded $|f(w)| \leq P$ for all $w \geq 0$ with a non-negative constant $P$.

A5 assumes that demand is non-negative, and that its density is supported on $[0, \infty)$, Lipschitz continuous, and bounded. This assumption is fairly mild and admits a large class of probability distributions, such as exponential, gamma, and truncated normal distributions.

We will use the above five assumptions for our theoretical analysis. Furthermore, we define the following function for convenience:

$$G(z,t) = cz + H(z,t) + \gamma E(J((z-w)^+,t+1))$$

Note that under the above definitions, we have the relation

$$\min_{u \in U_t} E\left[ g(x,t,u,w) + \gamma J(h(x,t,u,w),t+1) \right] = \min_{z \in [x,c_{\max}]} G(z,t) + K \cdot 1_{R_+}(z - x) - cx.$$
3.2. Proof Technique

To prove our main result, we will use proof by induction to show that the particular structure of the policy and value function is preserved. Our approach will be to show a number of results that hold under a temporary assumption (i.e., the induction hypothesis), which will then be proved to hold under the assumptions \textbf{A1–A5} in a final theorem that concludes our proof. Our temporary assumption is:

**Assumption T.** For any fixed \( t \in \{0, \ldots, k - 1\} \), we can represent \( J(z, t + 1) \) as

\[
J(z, t + 1) = J_i(z, t + 1) \text{ if } z \in Z_{i,t} \text{ for } i = 1, \ldots, n_t,
\]

where \( Z_{i,t} = [z_{i-1,t}, z_{i,t}] \) forms a partition of the domain \([0, c_{\text{max}}]\) with

\[
0 = z_{0,t} < z_{1,t} < \cdots < z_{n,t} = c_{\text{max}}.
\]

Moreover, the derivative of each piece \( \partial_z J_i(z, t + 1) \) is absolutely bounded by a finite, non-negative constant \( M_{t+1} \) and is Lipschitz continuous.

\textbf{AT} assumes that the value function is continuous and piecewise differentiable, and can also be partitioned into multiple pieces, the derivatives of which are Lipschitz and bounded. This assumption is required to characterize the policy structure that is complicated by the non-convexity induced by the reset control because these pieces represent the different regions of the optimal policy, which we investigate in the proceeding analysis.

3.3. Proof of the Optimal Policy Structure

To begin the proof, we first show a result on the smoothness of the cost-to-go function.

**Proposition 2.** If \textbf{A1–A5} and \textbf{AT} hold, then \( \mathbb{E}(J((z - w)^+, t + 1)) \) has a derivative that is Lipschitz continuous with non-negative constant \( (Lc_{\text{max}} + P)M_{t+1} \), and the derivative is absolutely bounded by \( M_{t+1} \).

The main proof of this proposition can be found in Appendix \( \text{A} \). Here, we present a brief sketch. First, we show that \( \mathbb{E}(J((z - w)^+, t + 1)) \) is differentiable using \textbf{AT} and the Leibniz integral rule for
derivatives. We then show that this derivative is Lipschitz continuous by introducing an auxiliary function that has the same expectation as the cost-to-go function. Integrating over the domain and using Lemma 2 gives this result. Finally, we show that the derivative is absolutely bounded by $M_{t+1}$ using standard integral and absolute value inequalities.

If the cost-to-go is piece-wise differentiable and the derivative of each piece is Lipschitz continuous, then the expected cost-to-go has a Lipschitz derivative. This is because taking the expectation smooths the function via convolution. Next, we show a result on the structure of $G(z,t)$.

**Proposition 3.** If $A1–A5$ and $AT$ hold, then the function $G(z,t)$ is continuous and convex on $z \in [0,c_{\text{max}}]$ for all fixed $\gamma$ such that $0 \leq \gamma \leq m_t/((Lc_{\text{max}} + P)M_{t+1})$.

The full proof of this proposition can be found in Appendix A, but here we present a sketch of the proof. To show this result, we note that $G$ is composed of a linear function, a twice differentiable function, and by Proposition 2, a function with a Lipschitz derivative. Since this implies that $G$ is absolutely continuous in its first argument, we demonstrate that it has a non-decreasing derivative, thus meaning it is convex on the desired interval.

This proposition intuitively states that the sum of the expected single stage cost (strongly convex) and the expected cost-to-go (non-convex but smooth) preserves convexity under suitable conditions, which can be computed using the parameters of the two functions. For example, $f_1(x) = x^2/2 + \gamma \sin(x)$ is convex for $\gamma \leq 1$, whereas $f_2(x) = x^2/2 - \gamma|x|$ will never be convex unless $\gamma = 0$. This exemplifies that the value of the discount factor is critical, but the smoothness of the expected cost-to-go, which was proved in Proposition 2, is even more crucial in retaining the convexity of the sum.

Our next result generalizes the known result (i.e., Lemma 4.2.1 in Bertsekas (1995)) in inventory management models for the case of a closed feasible set.

**Proposition 4.** If $G(\cdot,t)$ is continuous and convex on $[0,c_{\text{max}}]$, there exists $S$ such that $G(S,t) \leq G(z,t)$ for all $z \in [0,c_{\text{max}}]$. Furthermore, let $s = \inf \{z \in [0,S] \mid G(S,t) + K = G(z,t)\}$. If such $s$ exists, then
1. \( G(S, t) + K = G(s, t) \)

2. \( G(\cdot, t) \) is a non-increasing function on \([0, s]\)

3. \( G(S, t) + K \leq G(z, t) \) for all \( z \in [0, s] \)

4. \( G(y, t) \leq G(z, t) + K \) for all \( y, z \) with \( s \leq y \leq z \leq c_{\text{max}} \)

The full proof of this proposition can be found in Appendix A, but here we present a sketch of the proof. First, we note that because \( G(\cdot, t) \) is continuous on the closed interval, it must have a minimizer, which shows the existence of \( S \). Next, by continuity of \( G(\cdot, t) \), we note that there may exist \( s \) such that \( G(s, t) = G(S, t) + K \), thus proving \( C1 \). We then prove \( C3 \) using the convexity of \( G(\cdot, t) \). Then using these conditions and convexity, we prove \( C2 \). Finally, \( C4 \) follows from combining these results.

This proposition proves the optimality of the \((s, S)\) policy when the functions \( G(\cdot, t) \) are convex and have an upper and lower bound to the allowable values of the stock. We now extend this result to the functional structure present in the reset control problem.

**Proposition 5.** Suppose \( 0 \leq \gamma \leq m_t/(Lc_{\text{max}} + P)M_{t+1} \) for all \( t = 0, \ldots, k - 1 \). If \( A1-A5 \) and \( AT \) hold, then an optimal policy has the following four-stage structure with \((s_t, S_t, \sigma_t, \Sigma_t)\) thresholds:

1. If \( x_t \in [0, \sigma_t) \), then \( r^*_t = 1 \) and \( u^*_t = \varphi \)

2. If \( x_t \in [\sigma_t, s_t) \), then \( r^*_t = 0 \) and \( u^*_t = S_t - x_t \)

3. If \( x_t \in [s_t, \Sigma_t) \), then \( r^*_t = 0 \) and \( u^*_t = 0 \)

4. If \( x_t \geq \Sigma_t \), then \( r^*_t = 1 \) and \( u^*_t = \varphi \)

where \( \varphi \in [0, c_{\text{max}}] \) is a constant.

The full proof of this result can be found in Appendix A, but here we present a sketch. First, we define \( \varphi \) as the optimal value of \( u \) from the known reset state \( \zeta \) with \( t = 0 \). Because of this, from Proposition 1, we note that if the optimal reset action is \( r^*(x, t) = 1 \), that means \( u^*(x, t) = \varphi \), so we focus the proof on determining the reset policy. We now consider the set \( I_R \), which is the set of all states for which a reset action would not be taken. We then consider the case when this set is empty, meaning a reset action should be taken for all states. When this set is not empty,
define $\sigma_t, \sigma_t'$ as the infimum and supremum of $I_R$, respectively. Then by Proposition 4, we note that there exists $S_t' = \arg\min_{z \in [0,c_{max}]} G(z,t)$. However, we are not guaranteed for $s_t' = \sup\{z \in [0,S_t'] \mid G(S_t',t) + K \leq G(z,t)\}$ to exist. In the case where $s_t'$ does not exist, we demonstrate that the threshold $s_t$ should be set to $\sigma_t$ because $u^*(x,t) = 0$ will be the minimizer for all states in this case. Next, we show that if $s_t'$ does exist, then we should set the threshold $s_t$ to $s_t'$, since then the optimal action $u^*(x,t)$ will be to order up to level $S_t'$ if $x < s_t'$, and order nothing otherwise. Next, using continuity, convexity, Proposition 3 and A3, we show that the region of the state space where it is optimal not to take a reset action is the interval $[\sigma_t, \sigma_t'])$. This means that the optimal policy will use $\sigma_t$ value previously defined as one of the thresholds, and will use $\Sigma_t = \sigma_t'$ if $\sigma_t' < c_{max}$ and $\Sigma_t = +\infty$ otherwise. Next, we show that the threshold $S_t$ should be set to $S_t'$ by its definition and the properties of $s_t$ and $\sigma_t$.

The regions in Proposition 5 represent that (i) the inventory is sufficiently small such that reset cost is negligible, and thus resetting the system and ordering a one-step quantity is optimal; (ii) the stock is sufficient such that ordering up to a given quantity and consuming the majority prior to the stock becoming too obsolete or contaminated is optimal; (iii) the stock in the system is sufficient to satisfy possible future demands, and thus resetting the system or ordering additional stock is not optimal (a.k.a., do-nothing region); and (iv) there is so much stock in the inventory that it will almost certainly never be consumed, hence making it optimal to reset the system and reorder up to a baseline amount. Next, we show how this policy influences the structure of the cost-to-go function.

**Proposition 6.** Suppose $0 \leq \gamma \leq m_i / ((L c_{max} + P) M_{t+1})$ for all $t = 0, \ldots, k - 1$. If A1–A5 and A7 hold, then the function $J(x,t)$ for $t = 0, \ldots, k - 1$ has the following form with $(s_t, S_t, \sigma_t, \Sigma_t)$
thresholds:

\[
J(x,t) = \begin{cases} 
  J_0 + \mathbb{E}(s(x,t,w)), & x \in [0,\sigma_t) \\
  c(S_t - x) + K + H(S_t, t) + \gamma \mathbb{E}(J((S_t - w)^+, t+1)), & x \in [\sigma_t, s_t) \\
  H(x, t) + \gamma \mathbb{E}(J((x - w)^+, t+1)), & x \in [s_t, \Sigma_t) \\
  J_0 + \mathbb{E}(s(x,t,w)), & x \geq \Sigma_t 
\end{cases}
\] (28)

Furthermore, each piece has a Lipschitz derivative, and the derivative of each piece is absolutely bounded by \( M_t = \eta_t + c + \kappa_t + m_t/(L_{c,\text{max}} + P) \) for \( t = 0, \ldots, k - 1 \).

The full proof of this proposition can be found in Appendix A. Here, we present a proof sketch. The functional structure of the policy follows directly as a consequence of Proposition 5, so we focus on showing the properties of the derivative of the cost-to-go function. We thus examine each piece of the function individually. The result is proved for the first and fourth pieces as a consequence of A3, and we note that in the second region, \( J(x,t) \) is linear and thus also satisfies the result. For the third piece, the result follows as a consequence of Proposition 2 and A2.

The policy structure from Proposition 5 implies that the value function has up to four thresholds. The first two thresholds, namely \( s \) and \( S \), are obtained from the value function when there is no reset, which is linear and then convex. The points of intersection that result from juxtaposing the value functions with and without reset give us the last two thresholds, \( \sigma \) and \( \Sigma \), thereby yielding the optimal cost-to-go that is continuous and piecewise differentiable, as shown in Figure 1. We now complete the proof by induction and show that the optimal policy and cost-to-go functions indeed follow the forms above.

**Theorem 1.** Suppose \( \gamma \) is such that \( 0 \leq \gamma \leq \min \{ \gamma_t \mid t \in \{0, \ldots, k - 1\} \} \) for

\[
\gamma_{k-1} = \frac{m_{k-1}}{(L_{c,\text{max}} + P) \cdot \eta_k} \\
\gamma_t = \frac{m_t}{(L_{c,\text{max}} + P) \cdot (\eta_{t+1} + c + \kappa_{t+1}) + m_{t+1}} \quad \text{for } t = 0, \ldots, k - 2
\] (29)

If A1–A5 hold, then the policy described in Proposition 5 is optimal and the value function has the structure described in Proposition 6.
The complete proof of the theorem can be found in Appendix A, but here we present a sketch. First, we note that when \( t = k \), the optimal policy will choose to reset the system for all \( x \in [0, c_{\text{max}}] \), which means \( J(x, k) = J_0 + \mathbb{E}(s(x, k, w)) \). The rest follows by Propositions 5 and 6 and induction on \( t \).

This concludes our proof by induction, from which we obtain the threshold for the discount factor that guarantees the structural properties of the optimal policy that we proved using assumptions A1-A5. We can also safely remove the temporary assumption AT, as Proposition 6 proves that the value function can be partitioned into at most four pieces, each of which is Lipschitz continuous and absolutely bounded.

### 3.4. Policy Structure for Water Storage and Retail Management Problems

In this section, we provide sufficient conditions for both water storage and retail problems to show when their policies have the four-threshold structure of the reset problem. To do this, we show what conditions are sufficient such that each model satisfies assumptions A1-A5. First, we outline the conditions under which the water storage problem satisfies all assumptions.

**Proposition 7.** Suppose A5 (the assumption about the distribution of \( w_n \)) holds, and that \( f(w) > 0 \) for all \( w \in [0, c_{\text{max}}] \). If \( p - q(t) > 0 \) for all \( t = 0, \ldots, k - 1 \), then the water storage problem described in Section 2.1 satisfies A1-A5.
The main proof of this proposition can be found in Appendix A but here we present a sketch. First, we note that A1 and A4 hold by the definition of the problem, and that A5 holds by the assumptions of the proposition. Then we show that through a reformulation of the water problem stage cost (and setting the fixed cost to zero), the problem can be shown in conjunction with the Leibniz integral rule to satisfy A2. Next, we show that A3 follows from the structure of the reset conditions in this problem.

The following proposition presents sufficient conditions for the retail problem to have the same threshold policy structure.

**Proposition 8.** Suppose A5 holds, and that \( f(w) > 0 \) for all \( w \in [0, c_{\text{max}}] \). Then the retail management problem described in Section 2.2 satisfies A1–A5.

The main proof of this proposition can be found in Appendix A. However, we present a sketch here. First, we note that A1, A4, and A5 hold by definition of the problem and by assumption of the proposition. Next, letting the fixed cost \( K = k_u \) and using the Leibniz integral rule, the problem can be shown to satisfy A2. Finally, A3 follows from the reset cost structure.

4. Numerical Results

In this section, we present numerical studies to validate the theoretical structure we derived in Section 3. All computations are performed in MATLAB 2018b on a laptop computer with a 2.6GHz processor and 16GB of RAM. First, we present the numerical dynamic programming algorithm that we will use to compute the optimal value functions and policies. We then explore the two case studies of water storage and retail management problems and numerically verify that their value functions and optimal policies follow the theoretical structure.

4.1. Binary Dynamic Search Algorithm

We solve the dynamic programming equations and compute the optimal policy using the Binary Dynamic Search (BiDS) algorithm, initially developed by Mintz et al. (2017). The classical discounted reward settings, optimal value functions, and policies are computed with some form of value iteration (VI) and policy iteration (PI). However, if the state space is infinite, this computation
Algorithm 1 Binary Dynamic Search (BiDS) Algorithm (Mintz et al. 2017)

1: initialize $\underline{v} \leftarrow 0$ and $\overline{v} \leftarrow (\frac{1}{1-\gamma}) \min_u E[g(\zeta, 0, u, w_0) + \gamma \cdot s(h(\zeta, 0, u, w_0), 1, w_1)]$

2: repeat

3: set $v \leftarrow (\overline{v} + \underline{v}) / 2$

4: set $V(x, k, v) = v + E[s(x, w, k)]$

5: for $t = (k - 1), (k - 2), \ldots, 0$ do

6: set $V(x, t, v) = \min \{v + E[s(x, w, t)], \min_{u \in U_k} E[g(x, t, u, w) + \gamma V(h(x, t, u, w), t + 1, v)]\}$

7: end for

8: set $\Upsilon(v) = \min_{u \in U_0} E[g(\zeta, 0, u, w) + \gamma V(h(\zeta, 0, u, w), 1, v)]$

9: if $v > \Upsilon(v)$ then

10: set $\overline{v} \leftarrow v$

11: else

12: set $\underline{v} \leftarrow v$

13: end if

14: until $(\overline{v} - \underline{v}) \leq \epsilon$

15: set $v^* = (\overline{v} + \underline{v}) / 2$

is equivalent to finding a fixed point in an infinite-dimensional functional space, and while strong convergence guarantees exist in a discounted setting, the rate of convergence is highly dependent on the discount factor and can be practically slow (Bertsekas 1995). For these reasons, these problems are generally solved using approximate dynamic programming methods in practice. However, the BiDS algorithm can exploit the specific structure of reset control problems to calculate the optimal value function and policy with arbitrary precision. To do this, BiDS converts the problem into finding a fixed point in a vector space using binary search, where this vector can be thought of as the optimal value function evaluated at the reset state and time 0. As can be seen in Algorithm 1, BiDS first initializes the search space using an upper and lower bound on the value function at the reset state. Then BiDS takes the midpoint of the interval and uses backward induction to
compute the implied reset state cost. Depending on how this value compares with the candidate for the iteration, a new search interval is selected using the same procedures as a binary search. The algorithm terminates either when a true fixed point is found or when the numerical tolerance $\epsilon$ is reached. The theoretical convergence and computational guarantees of BiDS can be found in Mintz et al. (2017).

4.2. Example: Water Storage Problem

We use BiDS to numerically solve the water storage problem from Section 2.1. We generate random demand from the truncated normal distribution that is bounded from below by zero (Botev 2017) and set $k = 7$ to ensure that we flush at least once a week. The cost parameters are carefully chosen such that we place more weight on shortage cost, compared to purchasing and flushing costs (i.e., $p > c_u > c_r$). Furthermore, the cost of consuming water that has been stored long-term exceeds the cost of flushing water (i.e., $q(t) < p$ for all $t$, but $q(t) \geq c_r$ for $t$ close to $k$) such that the model favors resetting the tank more often to prevent contamination. In addition, while exponential microbial growth is normally assumed (Maier et al. 2009), we present a simplified model with linear growth to account for fluctuations in the microbial load that result from water usage and refilling. Note that any other health penalty could be used as long as it satisfies the monotonicity condition.

Figure 2  Example of an optimal policy for the water storage problem
In Figure 2, the top row shows the value function, and the second row shows the optimal control actions in blue (optimal fill amount) and red (reset control action). The bottom row shows the different zones for the optimal policy—yellow corresponds to flushing the tank and reordering water, cyan to ordering water and not flushing, and dark blue to not doing anything. For all plots, the state $x$ (the amount of water in the tank) is on the $x$-axis, and the subplots from left to right correspond to the states $t = 1, \ldots, 7$ (the number of days since the tank was last emptied).

With all other parameters fixed from Figure 2, we analyze the following two scenarios in Figure 3. When there is a high risk of contamination (top), we impose a larger penalty on water consumption to reflect this condition. We can observe that the optimal policy indeed suggests flushing the tank more often, as indicated by thicker yellow and thinner cyan bands, to reduce the risk of consuming contaminated water. Further exacerbation due to limited access to water (bottom), which is often the case in the developing world (United Nations 2015, WaterAid 2016), can be modeled by assigning a higher cost to purchasing water. Then we can see that the control policy is mostly composed of reset (yellow) and do-nothing (dark blue) regions, with their thresholds slightly shifted to the left compared to the previous policy. As illustrated by wider blue and narrower yellow regions, this regrettably implies that we cannot afford to reset that often. Nevertheless, our model provides a guideline that we can rely on, especially when we are faced with a difficult choice, whether it suggest a breakthrough or a compromise.
4.3. Example: Retail Management Problem

In this section, we solve the retail management problem described in Section 2.2 using the BiDS algorithm under two different cases, as shown in Figure 4. If a firm lacks the flexibility required to swiftly change the product line (top), we model this by assigning a higher reset cost. We can then observe that the optimal policy essentially reduces to an \((s, S)\) policy, similar to that of classical inventory management problems, where firms do not have the infrastructure and resources to switch the product line often. In contrast, if a company is situated in a rapidly evolving industry and thus equipped with the speed and adaptability demanded by the market (bottom), we use a lower reset cost but a higher holding cost to take the ephemeral nature of our merchandise into consideration.

We can now see that the inventory policy has a yellow region on the right. This encourages us to change inventory when there is too much on hand, and therefore epitomizes the right response to the transiency of a trend.

![Figure 4](image)

**Figure 4** Optimal policy for different cases of the retail management problem

At each \(t\), there are at most three thresholds that separate the control policy into four regions. In the first region (yellow), the trivial amount of stock renders the inventory reset cost negligible. Hence, emptying the inventory and ordering an optimal one-step quantity is optimal. In the next region (cyan), the sufficiency of stock in the inventory makes it optimal to order up to some quantity and still consume most of the stock before it becomes obsolete. In the third ‘do-nothing region’ (blue), it is not optimal to reset the inventory or order further stock as potential future demand can already be satisfied. In the final region (yellow), the inventory stock is almost certain to never be
consumed, and it is thus optimal to empty the inventory and reorder up to some optimal quantity, changing the product line if appropriate. Also, note that this covers all possible stages that can be suggested by the optimal policy, and that the actual policy may not have all of them, as in the case of Figure 4.

Figure 5  Example of an optimal policy for deteriorating item with linear time-dependent holding cost

5. Conclusion

Inventory management that empowers us to flexibly adapt to change is necessary to survive in a market and society that continues to evolve with growing acceleration. Previous inventory policies were contingent upon the tradeoff balance between holding and shortage costs, and informed us only of the optimal inventory position, or the quantity that must be ordered, for each period. By introducing the option to reset inventory into such classical frameworks, we are pioneering the design of more dynamic policies that overcome the limitations of pre-existing static policies. Furthermore, this enables us to continuously assess whether resets are more profitable at the preset periodic intervals or at earlier time points. In this paper, we (i) theoretically investigated the structural properties underlying the optimal policy of this problem, and (ii) implemented the appropriate algorithm to numerically solve and empirically validate the policy structure of the
aforementioned problem. Finally, we demonstrated the broad utility of this problem by providing both non-profit and for-profit examples.

The original motivation for this paper stemmed from the local water inventory management for developing nations (Mintz et al. 2017) as they often lack water distribution networks that provide clean water throughout the day. As a result, homes and apartments utilize water storage systems that are filled during a small window of time in the day when the water distribution network is active. However, these water storage systems are not equipped with disinfection capabilities. This leads to water being stored for long durations, thereby exposing people to substantially increased bacterial and viral water contamination (Coelho et al. 2003, Lee and Schwab 2005, Kumpel and Nelson 2016). Furthermore, obtaining even the minimal quantity of water necessary for survival can be a huge burden for people without a continuous water supply at home. This forces these individuals to compromise their health by using less water or collecting water from unsafe sources (WaterAid 2016). Our model seeks to promote universal and equitable access to safe water by providing guidance towards making decisions when confronted with such healthcare dilemmas.

Our formulation and proof are also sufficiently flexible to account for general reset control problems in supply chain management. For instance, we can incorporate modifications, such as introducing a time-varying holding cost to take the depreciation rate into account and penalize based on the amount of time it has been sitting on a shelf, with the purpose of managing aging inventories (see Figure 5). Finally, our results suggest that the sufficient conditions that ensure the threshold structure of the optimal policy are marginally conservative since they are not necessary conditions, and that there may exist a relaxation of these conditions, which we leave for future studies.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant CMMI-1847666.
Appendix A: Proof of Propositions and Theorem

**Proof of Proposition 3.** We first show that \( E((z-w)^+, t+1) \) is differentiable. Consider any \( z \in [0, c_{\text{max}}] \), and observe that

\[
J((z-w)^+, t+1) = \begin{cases} 
J_1(0, t+1), & \text{if } z-w \leq 0 \\
J_i(z-w, t+1), & \text{if } z-w \in Z_{i,t}
\end{cases}
\]

Now, let \( m \) be the smallest integer such that \( z \in Z_{m,t} \). Then the expectation is given by

\[
E(J((z-w)^+, t+1)) = \int_0^\infty J((z-w)^+, t+1) f(w) dw = \int_0^{z-z_{m-1},t} J_m(z-w, t+1) f(w) dw + \\
\sum_{i=1}^{m-1} \int_{z-z_{i-1},t}^{z-z_i,t} J_i(z-w, t+1) f(w) dw + \int_z^\infty J_i(0, t+1) f(w) dw.
\]

(A.2)

Since by assumption both \( J_i(z, t+1) \) and \( \partial_z J_i(z, t+1) \) are continuous for \( z \in Z_{i,t} \), using the Leibniz integral rule gives

\[
\partial_z E(J((z-w)^+, t+1)) = J_m(z_{m-1, t}, t+1) f(z-z_{m-1, t}) + \int_0^{z-z_{m-1},t} \partial_z J_m(z-w, t+1) f(w) dw + \\
\sum_{i=1}^{m-1} \left( J_i(z_{i-1, t}, t+1) f(z-z_{i-1, t+1}) - J_i(z_{i,t}, t+1) f(z-z_{i,t}) + \int_{z-z_{i-1},t}^{z-z_i,t} \partial_z J_i(z-w, t+1) f(w) dw \right) - \\
\int_z^\infty \partial_z J_1(0, t+1) f(w) dw = \int_0^{z-z_{m-1},t} \partial_z J_m(z-w, t+1) f(w) dw + \\
\sum_{i=1}^{m-1} \int_{z-z_{i-1},t}^{z-z_i,t} \partial_z J_i(z-w, t+1) f(w) dw + \int_z^\infty \partial_z J_1(0, t+1) f(w) dw = \\
\int_0^\infty \partial_z J((z-w)^+, t+1) f(w) dw
\]

(A.3)

where the last equality follows from the fact that \( f(\cdot) \) is Lipschitz continuous and \( J(\cdot, t+1) \) is differentiable almost everywhere on its domain since it is piecewise differentiable by assumption. Hence, we can conclude that \( E(J((z-w)^+, t+1)) \) is differentiable.

Next, we show that this derivative is Lipschitz continuous. Define

\[
J_A(x, t+1) = \begin{cases} 
J(0, t+1), & x \leq 0 \\
J_0(x, t+1), & x \in [0, c_{\text{max}}] \\
J(c_{\text{max}}, t+1), & x \geq c_{\text{max}}
\end{cases}
\]

(A.4)
where the subscript \( A \) indicates “auxiliary”. Then for \( z \in [0, c_{\text{max}}] \) and \( w \in [0, \infty) \), we have \( J((z - w)^+, t + 1) = J_A(z - w, t + 1) \). Note that \( \mathbb{E}(J_A(z - w, t + 1)) = \int_{0}^{\infty} J_A(z - w, t + 1) f(w) dw = \int_{-\infty}^{\infty} J_A(z - w, t + 1) f(w) dw = (J_A(\cdot, t) * f)(z) \) (i.e., the convolution of \( J_A(\cdot, t + 1) \) and \( f(\cdot) \)), where the second equality follows since \( f(w) = 0 \) for \( w \in (-\infty, 0] \) by assumption. This implies that

\[
\partial_z (J_A(\cdot, t + 1) * f)(z) = \partial_z \mathbb{E}(J_A(z - w, t + 1))
\]

\[
= \partial_z \mathbb{E}(J((z - w)^+, t + 1))
\]

\[
= \int_{0}^{\infty} \partial_z J((z - w)^+, t + 1) f(w) dw
\]

\[
= \int_{0}^{\infty} \partial_z J_A(z - w, t + 1) f(w) dw
\]

\[
= (\partial_z J_A(\cdot, t + 1) * f)(z)
\]

Observe that

\[
\partial_z J_A(z, t + 1) = \begin{cases} 
0, & x < 0 \\
\partial_z J(z, t + 1), & z \in (0, c_{\text{max}}) \\
0, & x > c_{\text{max}}
\end{cases}
\]

(A.6)

This means \( \int_{-\infty}^{\infty} |\partial_z J_A(z, t + 1)| dz = \int_{0}^{c_{\text{max}}} |\partial_z J(z, t + 1)| dz \leq c_{\text{max}} M_{t+1} \) since we had assumed \( |\partial_z J(z, t + 1)| \leq M_{t+1} \). Thus, Lemma 2 implies \( \partial_z \mathbb{E}(J((z - w)^+, t + 1)) \) is Lipschitz with constant \( (L c_{\text{max}} + P) M_{t+1} \).

Lastly, we show that the derivative is absolutely bounded by \( M_{t+1} \). From (A.5), we have

\[
|\partial_z (J_A(\cdot, t + 1) * f)(z)| = |(\partial_z J_A(x, t + 1) * f)(z)|
\]

\[
= \left| \int_{0}^{\infty} \partial_z J((z - w)^+, t + 1) f(w) dw \right|
\]

\[
\leq \int_{0}^{\infty} |\partial_z J((z - w)^+, t + 1)| \cdot |f(w)| dw
\]

\[
\leq M_{t+1} \int_{0}^{\infty} |f(w)| dw
\]

\[
\leq M_{t+1}
\]

where in the last line we have used the facts that \( f(w) \) is a probability density, and hence non-negative, and integrates to one. \( \square \)

**Proof of Proposition 5** We first recall the definition

\[
G(z, t) = c \cdot z + H(z, t) + \gamma \mathbb{E}(J((z - w)^+, t + 1)).
\]

Note that \( c \cdot z \) is linear, that \( H(z, t) \) is twice differentiable by assumption, and that \( \mathbb{E}(J((z - w)^+, t + 1)) \) has a Lipschitz derivative by Proposition 2. Hence, \( G(\cdot, t) \) is absolutely continuous in its first argument. Since the
domain \([0, \max]\) is closed and bounded, this means that \(G(\cdot, t)\) is convex if it has a non-decreasing derivative (see page 115 of Fremlin (2001)). From A2, we have that \(\partial_z (c \cdot z + H(z, t)) \geq m_t\), and Proposition 2 gives that \(\partial_z (J((z - w)^+, t + 1))\) is Lipschitz with constant \((Lc_{\max} + P)M_{t+1}\). Thus, by Lemma \([\text{A1}]\) we have that \(\partial_z G(z, t)\) is non-decreasing for \(\gamma\) such that \(0 \leq \gamma \leq m_t / ((Lc_{\max} + P)M_{t+1})\). □

**Proof of Proposition \([\text{A3}]\)** Since \(G(\cdot, t)\) is continuous on the closed interval \([0, \max]\), there exists a minimizer of \(G(\cdot, t)\). Let \(S = \arg\min_{x \in [0, C_{\max}]} G(z, t)\). For the remainder of the proof, we consider the case where \(S\) exists. By the continuity of \(G(\cdot, t)\) and the definition of \(S\), we must have \(G(S, t) + K = G(s, t)\). Next, for all \(z \in [0, S]\), there exists \(\lambda \in [0, 1]\) such that \(S = \lambda z + (1 - \lambda)S\) and

\[
G(s, t) \leq \lambda G(z, t) + (1 - \lambda)G(S, t) \leq \lambda G(z, t) + (1 - \lambda)G(z, t) = G(z, t),
\]

(A.8)

where the inequalities follow from the convexity of \(G(\cdot, t)\) and the definition of \(S\), respectively. Next, consider any \(z_1\) and \(z_2\) with \(0 \leq z_1 \leq z_2 \leq s\), and note that there exists \(\lambda \in [0, 1]\) such that \(z_2 = \lambda z_1 + (1 - \lambda)s\). Then we have

\[
G(z_2, t) \leq \lambda G(z_1, t) + (1 - \lambda)G(s, t) \leq \lambda G(z_1, t) + (1 - \lambda)G(z_1, t) = G(z_1, t),
\]

(A.9)

where the second inequality follows by (A.8). This shows \(G(\cdot, t)\) is non-increasing on \([0, S]\). Since \(G(\cdot, t)\) is continuous, this means by definition of \(S\) that \(G(S, t) + K \leq G(s, t)\). Using the above result that \(G(\cdot, t)\) is non-increasing on \([0, S]\), we have \(G_t(S) + K \leq G_t(z)\) for all \(z \in [0, S]\). By definition of \(S\), for any \(y \in [S, \max]\), we have \(G_t(S) + K \geq G_t(y)\). However, we also know that \(G_t(S) \leq G_t(z)\) for any \(z \in [y, \max]\). Combining the two yields the last part of the result. □

**Proof of Proposition \([\text{A4}]\)** We first define the constant

\[
\varphi = \arg\min_{w \in \mathcal{E}} \mathbb{E} \left[ g(\zeta, 0, u, w) + \gamma J(h(\zeta, 0, u, w), 1) \right].
\]

(A.10)

From Proposition 1, we have \(u_0^*(x) = \varphi\) whenever \(r_0^*(x) = 1\). So, our proof will be structured around determining an optimal reset policy \(r_0^*(x)\). We define the functions

\[
J_R(x, t) = J_0 + \mathbb{E} (s(x, t, w))
\]

\[
J_N(x, t) = \min_{u \in [0, \max]} G(x + u, t) + K \cdot 1_{x < z_1}(u) - c \cdot x.
\]

(A.11)

where the subscript \(R\) indicates “reset”, and the subscript \(N\) indicates “no reset”. Observe that these are the value functions corresponding to \(r_t = 1\) and \(r_t = 0\), respectively. Next, define the set \(I_R = \{ x \in [0, \max] : J_R(x, t) \geq J_N(x, t) \}\), and observe that \(I_R\) is bounded by construction. Now, we consider two cases:
The first case is when $I_R = \emptyset$. Then by definition of $I_R$, we have $J_R(x, t) < J_N(x, t)$ for all $x \in [0, c_{\text{max}}]$. For this case, an optimal policy is to choose $\sigma_t = s_t = \Sigma_t = +\infty$ since it is optimal to choose $r^*_t(x) = 1$ for all $x \in [0, c_{\text{max}}]$.

The second case is when $I_R \neq \emptyset$. Let $\sigma_t = \inf\{x \in [0, c_{\text{max}}] \mid J_R(x, t) \geq J_N(x, t)\}$ and $\sigma'_t = \sup\{x \in [0, c_{\text{max}}] \mid J_R(x, t) \geq J_N(x, t)\}$, and note that $\sigma_t$ and $\sigma'_t$ are finite since $I_R$ is bounded and non-empty. Let $S'_t = \arg\min_{x \in [0, c_{\text{max}}]} G(z, t)$ and define $s'_t = \sup\{z \in [0, S'_t] \mid G(S'_t, t) + K \leq G(z, t)\}$. Note that by Proposition 4, $S'_t$ is guaranteed to exist whereas $s'_t$ may or may not exist. We consider two subcases based on the existence of $s'_t$:

The first subcase is when $s'_t$ does not exist. Then an optimal policy chooses $s_t = \sigma_t$ because in this subcase, a minimizer to the optimization problem defining $J_N(x, t)$ is $u^*_t(x) = 0$ for all $x \in [0, c_{\text{max}}]$. Note that by setting $s_t$ equal to $\sigma_t$, the second policy region vanishes, which ensures that $u^*_t(x) = 0$ for all $x \in [0, c_{\text{max}}]$ with our policy.

The rest of the proof considers the second subcase in which $s'_t$ exists. We set $s_t = s'_t$, since

$$u^*_t(x) = \begin{cases} S'_t - x, & \text{if } 0 \leq x < s'_t \\ 0, & \text{if } s'_t \leq x \leq c_{\text{max}} \end{cases} \tag{A.12}$$

is optimal for the optimization problem defining $J_N(x, t)$. We now observe that for $x \in [0, s_t)$, $J_N(x, t) = G(S'_t, t) + K - c \cdot x$, which is non-increasing in $x$. For $x \in [s_t, c_{\text{max}}]$, $J_N(x, t) = G(x, t) - c \cdot x$, which is convex in $x$ since $G(x, t)$ is convex by Proposition 3. Clearly, $J_N(x, t)$ is continuous for $x \in [0, s_t)$ since it is linear in this region, and $J_N(x, t)$ is continuous for $x \in (s_t, c_{\text{max}}]$ since $G(\cdot, t)$ is continuous by Proposition 3. The only question about continuity occurs at $x = s_t$. Since $G(S'_t, t) + K = G(s_t, t)$ by definition, we have that $G(S'_t, t) + K - c \cdot s_t = G(s_t, t) - c \cdot s_t$. This proves that $J_N(x, t)$ must be continuous at $x = s_t$ since the left and right side of the last equality with $G(\cdot, t)$ are the limits of $J_N(\cdot)$ in the two respective regions. This means $J_N(x, t)$ is continuous on $x \in [0, c_{\text{max}}]$. Because $J_R(\cdot, t)$ is continuous by A3, this means that $J_R(\sigma_t, t) \geq J_N(\sigma_t, t)$ and $J_R(\sigma'_t, t) \geq J_N(\sigma'_t, t)$ by the definitions of $\sigma_t$ and $\sigma'_t$.

Next, consider any $x \in [\sigma_t, s_t]$ (note that our argument still holds even if this set is empty). Since $J_N(x, t)$ is non-increasing for $x \in [0, s_t]$, and since $J_R(x, t)$ is non-decreasing in $x$ by A3, we have

$$J_R(x, t) \geq J_R(\sigma_t, t) \geq J_N(\sigma_t, t) \geq J_N(x, t) \tag{A.13}$$
for \( x \in [\sigma_t, s_t] \). This argument also implies that \( J_R(s_t, t) \geq J_N(s_t, t) \). Next, consider any \( x \in [s_t, \sigma_t'] \), and observe that there exists \( \mu \in [0, 1] \) such that \( x = \mu s_t + (1 - \mu)\sigma_t' \). Since we showed above that \( J_N(\cdot, t) \) is convex, this means that for any \( x \in [s_t, \sigma_t'] \), we have

\[
J_N(x, t) = J_N(\mu s_t + (1 - \mu)\sigma_t', t) \leq \mu J_N(s_t, t) + (1 - \mu)J_N(\sigma_t', t)
\]

\[\leq \mu J_R(s_t, t) + (1 - \mu)J_R(\sigma_t', t) \tag{A.14}\]

\[\leq J_R(\mu s_t + (1 - \mu)\sigma_t', t) = J_R(x, t), \]

where the last inequality follows because \( J_R(\cdot, t) \) is concave by A3. Combining the above shows that \( J_R(x, t) \geq J_N(x, t) \) for all \( x \in [\sigma_t, \sigma_t'] \).

Now, observe that \( J_R(x, t) < J_N(x, t) \) for \( x \in [0, \sigma_t] \) and for \( x \in (\sigma_t', c_{\max}] \). (If this last statement were not true, then we could choose an \( x' \in [0, \sigma_t] \) or an \( x' \in (\sigma_t', c_{\max}] \) such that \( J_R(x', t) \geq J_N(x', t) \). Hence, we would have \( x' \in I_R \), which reaches a contradiction since, by the definition of \( \sigma_t \) and \( \sigma_t' \), we would have \( \sigma_t \leq x' \) or \( \sigma_t' \geq x' \).) Consequently, an optimal policy uses this value of \( \sigma_t \). If \( \sigma_t' < c_{\max} \), then an optimal policy uses \( \Sigma_t = \sigma_t' \), and if \( \sigma_t' = c_{\max} \), then an optimal policy chooses \( \Sigma_t = +\infty \).

Finally, we must choose a correct value for \( S_t \). We must consider three sub-subcases. The first sub-subcase is when \( S_t' < \sigma_t \). Since \( s_t' \leq S_t' \) by definition of \( s_t' \), this means the second policy region is empty when \( s_t = s_t' \), which ensures that \( u_t^*(x) = 0 \) for all \( x \in [\sigma_t, \sigma_t'] \) with our policy. Thus, we can choose \( S_t = S_t' \). (Any arbitrary choice of \( S_t \) would give an optimal policy because the corresponding region is empty.) The second sub-subcase is when \( \sigma_t \leq S_t' \leq \sigma_t' \). Then an optimal policy chooses \( S_t = S_t' \). The third sub-subcase is when \( S_t' > \sigma_t' \). By definition of \( \sigma_t \) and \( \sigma_t' \), this means \( J_R(S_t', t) < J_N(S_t', t) \) and \( \sigma_t' \geq \sigma_t \). Thus, in this sub-subcase, we have

\[
J_N(\sigma_t', t) = G(\sigma_t', t) - c \cdot \sigma_t' \geq G(S_t', t) - c \cdot S_t' = J_N(S_t', t), \tag{A.15}\]

where we have used the definition of \( S_t' \) in the last inequality. Recalling that \( J_R(x, t) \geq J_N(x, t) \) for all \( x \in [\sigma_t, \sigma_t'] \), we have

\[
J_R(S_t', t) < J_N(S_t', t) \leq J_N(\sigma_t', t) \leq J_R(\sigma_t', t). \tag{A.16}\]

However, this last statement is a contradiction since \( \sigma_t' < S_t' \) and \( J_R(\cdot, t) \) is non-decreasing by A3. Therefore, this sub-subcase is not possible. □

**Proof of Proposition 5** The policy structure in Proposition 5 implies that the value function \( J(x, t) \) takes the form \( [28] \). This means we have to analyze the derivative of at most four pieces. In the first and
fourth regions, by A3, we have that \( J(x,t) \) has a derivative that is Lipschitz with constant \( L_t \), and that the derivative is absolutely bounded by \( \eta_t \). In the second region, \( J(x,t) \) is linear in \( x \) and thus has a Lipschitz derivative that is absolutely bounded by \( c \). In the third region, we know that the first term has a Lipschitz derivative because it is twice differentiable by A2, and that the second term has a Lipschitz derivative by Proposition 2. Thus, the third region has a Lipschitz derivative because it is the sum of two functions with Lipschitz derivatives. To bound its derivative, we note that \( \partial_t H(x,t) \) is absolutely bounded by \( \kappa_t \) by A2.

From Proposition 2, we know that the derivative of the second term is absolutely bounded by \( \gamma M_t+1 \). Thus, the third piece has a derivative that is absolutely bounded by \( \kappa_t + \gamma M_t+1 \). This means the derivative of each piece is absolutely bounded by \( \eta_t + c + \kappa_t + \gamma M_t+1 \). As \( 0 \leq \gamma \leq m_t/((Lc_{\text{max}} + P)M_t+1) \), the derivative of each piece is absolutely bounded by \( M_t = \eta_t + c + \kappa_t + m_t/(Lc_{\text{max}} + P) \). □

**Proof of Theorem 3** First, note that we can choose \( M_k = \eta_k \) using A3, since an optimal policy at \( t = k \) is to choose \( r_k^*(x) = 1 \) for all \( x \in [0,c_{\text{max}}] \), which means \( J(x,k) = J_0 + E(s(x,k,w)) \). Inductively applying Propositions 5 and 6 implies the desired result. □

**Proof of Proposition 5** A1 and A4 hold by the definition of the water problem, and A5 holds by assumption. To show A2, we first note that

\[
\xi_n + u_n - (\xi_n + u_n - w_n)^+ = \xi_n + u_n - (\xi_n + u_n - w_n)^+ - (\xi_n + u_n - w_n)^- + (\xi_n + u_n - w_n)^-
\]

\[
= \xi_n + u_n - (\xi_n + u_n - w_n) + (\xi_n + u_n - w_n)^-
\]

\[
= w_n + (\xi_n + u_n - w_n)^-
\]

This means the single stage cost can be rewritten as

\[
g(\xi_n,\tau_n,u_n,w_n) = c_u u_n - p \cdot (\xi_n + u_n - w_n)^- + q(\tau_n) \cdot (\xi_n + u_n - (\xi_n + u_n - w_n)^+)
\]

\[
= c_u u_n - (p - q(\tau_n)) \cdot (\xi_n + u_n - w_n)^- + q(\tau_n) \cdot w_n
\]
Hence, (24) is satisfied by setting $K = 0$ and

$$H(z, t) = -(p - q(t)) \int_{z}^{\infty} (z - w)f(w)dw + q(t) \cdot \mathbb{E}(w). \quad (A.19)$$

Note that using the Leibniz integral rule twice gives

$$\partial_z H(z, t) = -(p - q(t)) \int_{z}^{\infty} f(w)dw$$

$$\partial_{zz} H(z, t) = (p - q(t)) \cdot f(z) \quad (A.20)$$

Thus, we have $|\partial_z H(z, t)| = |-(p - q(t)) \int_{z}^{\infty} f(w)dw| \leq p - q(t)$ since $f(w)$ is a density. Since $H(\cdot, t)$ is twice differentiable, this means it is absolutely continuous. Thus, $H(\cdot, t)$ is strongly convex on $[0, c_{\text{max}}]$ since $\partial_{zz} H(z, t) = (p - q(t)) \cdot f(z) \geq (p - q(t)) \cdot \inf_{z \in [0, c_{\text{max}}]} f(z) = (p - q(t)) \cdot \min_{z \in [0, c_{\text{max}}]} f(z) > 0$ for $z \in [0, c_{\text{max}}]$, where the second equality holds because a Lipschitz continuous function attains its minimum in a compact domain. This shows that A2 holds. To show that A3 holds, we first note that the water storage problem as described in Section 2.1 implies $\zeta = 0$. Next, note that the reset cost $R(x, t) = c_r x$ is linear. Hence, the conditions of A3 follow immediately. \( \Box \)

**Proof of Proposition 8** A1 and A4 hold by the definition of the retail problem, and A5 holds by assumption. To show A2, we first note that the single stage cost can be written as

$$g(\xi_n, r_n, u_n, w_n) = c_u u_n + k_u \cdot 1_{\mathbb{R}_{\geq 0}}(u_n) - p \cdot (\xi_n + u_n - w_n)^- + q \cdot (\xi_n + u_n - w_n)^+ \quad (A.21)$$

Hence, (24) is satisfied by setting $K = k_u$ and

$$H(z, t) = -p \int_{z}^{\infty} (z - w)f(w)dw + q \int_{-\infty}^{z} (z - w)f(w)dw. \quad (A.22)$$

Note that using the Leibniz integral rule twice gives

$$\partial_z H(z, t) = -p \int_{z}^{\infty} f(w)dw + q \int_{-\infty}^{z} f(w)dw$$

$$\partial_{zz} H(z, t) = (p + q) \cdot f(z) \quad (A.23)$$

Thus, we have $|\partial_z H(z, t)| = |p \int_{z}^{\infty} f(w)dw + q \int_{-\infty}^{z} f(w)dw| = (p + q) \cdot F(z) - p \leq q$, since $F(z) \leq 1$ for $z \in [0, c_{\text{max}}]$. Since $H(\cdot, t)$ is twice differentiable, this means it is absolutely continuous. Thus, $H(\cdot, t)$ is strongly convex on $[0, c_{\text{max}}]$ since $\partial_{zz} H(z, t) = (p + q) \cdot f(z) \geq (p + q) \cdot \min_{z \in [0, c_{\text{max}}]} f(z) > 0$ for $z \in [0, c_{\text{max}}]$. This shows that A2 holds. To show A3 holds, we first note that the retail management problem as described in Section 2.2 implies $\zeta = 0$. Next, note that the reset cost $R(x, t) = c_r x + k_r$ is linear. Hence, the conditions of A3 follow immediately. \( \Box \)
Appendix B: Statement and Proof of Lemmas

**Lemma 1.** Consider two functions $f(x)$ and $g(x)$ defined on a compact interval $[a, b]$. Suppose that $f(x)$ is differentiable, that its derivative is bounded from below by a positive constant $D_x f(x) \geq M > 0$, and that $g(x)$ is Lipschitz continuous with constant $L \geq 0$. Then for any fixed $\lambda$ such that $0 \leq \lambda \leq M/L$, the function $f(x) + \lambda g(x)$ is non-decreasing and continuous.

**Proof of Lemma 1.** The continuity of $f(x) + \lambda g(x)$ is immediate by the differentiability of $f(x)$ and the Lipschitz continuity of $g(x)$. So we focus on showing $f(x) + \lambda g(x)$ is non-decreasing. Consider any $x_1, x_2 \in [a, b]$ with $x_2 \geq x_1$. Since $D_x f(x) \geq M$, we have $f(x_2) - f(x_1) \geq M(x_2 - x_1)$. Since $g(x)$ is Lipschitz continuous with constant $L$, we have $|g(x_2) - g(x_1)| \leq L(x_2 - x_1)$. Combining these two inequalities implies

$$f(x_2) - f(x_1) + \lambda(g(x_2) - g(x_1)) \geq M(x_2 - x_1) - \lambda L(x_2 - x_1) \geq (M - \lambda L) \cdot (x_2 - x_1) \geq 0.$$  

(B.1)

where the last inequality follows because $x_2 \geq x_1$ and because $0 \leq \lambda \leq M/L$, which means $M - \lambda L \geq 0$. Since we have shown $f(x_2) - f(x_1) + \lambda(g(x_2) - g(x_1)) \geq 0$ for any $x_2 \geq x_1$, this implies $f(x) + \lambda g(x)$ is non-decreasing. □

**Lemma 2.** If $\int_{-\infty}^{\infty} |f(x)| \, dx$ is finite, $\int_{-\infty}^{\infty} |g(x)| \, dx$ is finite, $f(x) = 0$ for $x < 0$, $|f(x)| \leq M$, $|g(x)| \leq P$, and on $x \geq 0$, we have that $f(x)$ is Lipschitz with constant $L$, then the convolution $(f * g)(x) = \int_{-\infty}^{\infty} f(x-z)g(z) \, dz$ is Lipschitz with constant $(L \cdot \int_{-\infty}^{\infty} |g(z)| \, dz + MP)$.

**Proof of Lemma 2.** Without loss of generality, we assume that $x \leq y$. Observe that

$$|(f * g)(x) - (f * g)(y)| = |\int_{-\infty}^{\infty} f(x-z)g(z) \, dz - \int_{-\infty}^{\infty} f(y-z)g(z) \, dz|$$

$$= |\int_{-\infty}^{x} (f(x-z) - f(y-z)) \cdot g(z) \, dz - \int_{x}^{y} f(y-z) \cdot g(z) \, dz|$$

$$\leq \int_{-\infty}^{x} |f(x-z) - f(y-z)| \cdot |g(z)| \, dz + \int_{x}^{y} |f(y-z)| \cdot |g(z)| \, dz$$

$$\leq \int_{-\infty}^{x} L |(x-z) - (y-z)| \cdot |g(z)| \, dz + \int_{x}^{y} MP \cdot |g(z)| \, dz$$

$$\leq \int_{-\infty}^{x} L |x-y| \cdot |g(z)| \, dz + MP \cdot |x-y|$$

$$\leq (L \cdot \int_{-\infty}^{\infty} |g(z)| \, dz + MP) \cdot |x-y|$$

This shows the convolution is Lipschitz continuous. □
Lemma 3. Consider a function $f(x)$ that is finite at the point $y$ (i.e., $|f(y)|$ is bounded). If $f(x)$ is Lipschitz with constant $L$ on a compact domain $[a, b]$ with $y \in [a, b]$, then it is finitely bounded $|f(x)| \leq L(b-a) + |f(y)|$ for $x \in [a, b]$.

Proof of Lemma 3. For any $x \in [a, b]$, we have

$$
|f(x)| = |f(x) - f(y) + f(y)|
$$

$$
\leq |f(x) - f(y)| + |f(y)|
$$

$$
\leq L|x - y| + |f(y)|
$$

$$
\leq L(b-a) + |f(y)|
$$

(B.3)

Since the interval $[a, b]$ is compact, this means $a, b$ are finite. This gives the desired bound. □

References

Arrow KJ, Harris T, Marschak J (1951) Optimal inventory policy. *Econometrica* 19(3):250–272.

Barnes L, Lea-Greenwood G (2006) Fast fashioning the supply chain: shaping the research agenda. *Journal of Fashion Marketing and Management* 10(3):259–271.

Bertsekas DP (1995) *Dynamic programming and optimal control*, volume 1 (Athena Scientific).

Bhardwaj V, Fairhurst A (2010) Fast fashion: response to changes in the fashion industry. *The International Review of Retail, Distribution and Consumer Research* 20(1):165–173.

Blackburn J, Scudder G (2009) Supply chain strategies for perishable products: the case of fresh produce. *Production and Operations Management* 18(2):129–37.

Botev ZI (2017) The normal law under linear restrictions: simulation and estimation via minimax tilting. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79(1):125–148.

Cachon GP, Swinney R (2011) The value of fast fashion: Quick response, enhanced design, and strategic consumer behavior. *Management Science* 57(4):778–795.

Caro F, Gallien J (2007) Dynamic assortment with demand learning for seasonal consumer goods. *Management Science* 53(2):276–292.

Caro F, Gallien J (2010) Inventory management of a fast-fashion retail network. *Operations Research* 58(2):257–273.
Caro F, Martínez-de-Albéniz V (2015) Fast fashion: Business model overview and research opportunities. Agrawal N, Smith SA, eds., Retail Supply Chain Management, 237–264, International Series in Operations Research & Management Science (Springer).

Chu LY, Hsu VN, Shen Z (2005) An economic lot-sizing problem with perishable inventory and economies of scale costs: Approximation solutions and worst case analysis. Naval Research Logistics 52(6):536–548.

Coelho LC, Laporte G (2014) Optimal joint replenishment, delivery and inventory management policies for perishable products. Computers & Operations Research 47:42–52.

Coelho ST, James S, Sumna N, Abu Jaish A, Chatila J (2003) Controlling water quality in intermittent supply systems. Water Supply 3(1-2):119–125.

de Farias DP, Van Roy B (2003) The linear programming approach to approximate dynamic programming. Operations Research 51(6):850–865.

Farahani P, Grunow M, Günther HO (2012) Integrated production and distribution planning for perishable food products. Flexible Services and Manufacturing Journal 24:28–51.

Fremlin DH (2001) Measure Theory, volume 2 (Torres Fremlin).

Gadgil A (1998) Drinking water in developing countries. Annual Review of Energy and the Environment 23(1):253–286.

Hadley G, Whitin T (1963) Analysis of inventory systems (Prentice-Hall, Inc.).

Haskell WB, Jain R, Kalathil D (2016) Empirical dynamic programming. Mathematics of Operations Research 41(2):402–429.

Hsu VN (2000) Dynamic economic lot size model with perishable inventory. Management Science 46(8):1159–1169.

Hsu VN, Lowe TJ (2001) Dynamic economic lot size models with period-pair-dependent backorder and inventory costs. Operations Research 49(2):316–321.

Karaesmen IZ, Scheller-Wolf A, Deniz B (2011) Managing perishable and aging inventories: Review and future research directions. Kempf KG, Keskinocak P, Uzsoy R, eds., Planning Production and Inventories in the Extended Enterprise: A State of the Art Handbook, 393–436 (Springer).
Kariotoglou N, Summers S, Summers T, Kamgarpour M, Lygeros J (2013) Approximate dynamic programming for stochastic reachability. *2013 European Control Conference (ECC)*, 584–589.

Klingel P (2012) Technical causes and impacts of intermittent water distribution. *Water Supply* 12(4):504–512.

Kumpel E, Nelson KL (2016) Intermittent water supply: Prevalence, practice, and microbial water quality. *Environmental Science & Technology* 50(2):542–553.

Lee EJ, Schwab KJ (2005) Deficiencies in drinking water distribution systems in developing countries. *Journal of Water and Health* 3(2):109–127.

Maier RM, Pepper IL, Gerba CP (2009) *Environmental Microbiology* (Elsevier).

Matsuo H, Ogawa S (2007) Innovating innovation: The case of seven-eleven japan. *International Commerce Review* 7(2).

Mintz Y, Shen Z, Aswani A (2017) Local water storage control for the developing world. *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 5074–5079.

Nagurney A, Yu M, Masoumi AH, Nagurney LS (2013) *Networks against time: Supply chain analytics for perishable products* (Springer).

Nahmias S (1982) Perishable inventory theory: A review. *Operations Research* 30(4):680–708.

Pierskalla WP (2005) Supply chain management of blood banks. *Operations research and health care*, 103–145 (Springer).

Pierskalla WP, Roach CD (1972) Optimal issuing policies for perishable inventory. *Management Science* 18(11):603–614.

Powell WB (2007) *Approximate Dynamic Programming: Solving the curses of dimensionality* (John Wiley & Sons).

Prastacos GP (1984) Blood inventory management: an overview of theory and practice. *Management Science* 30(7):777–800.

Rosenfield DB (1989) Disposal of excess inventory. *Operations Research* 37(3):404–409.
Rosenfield DB (1992) Optimality of myopic policies in disposing excess inventory. *Operations Research* 40(4):800–803.

Ryzhov IO, Powell WB, Frazier PI (2012) The knowledge gradient algorithm for a general class of online learning problems. *Operations Research* 60(1):180–195.

Scarf H (1959) The optimality of (S, s) policies in the dynamic inventory problem. Technical report, Stanford University.

Shen Z, Dessouky M, Ordóñez F (2011) Perishable inventory management system with a minimum volume constraint. *Journal of the Operational Research Society* 62(12):2063–2082.

Snyder LV, Shen Z (2019) *Fundamentals of Supply Chain Theory* (Wiley).

Tokajian S, Hashwa F (2003) Water quality problems associated with intermittent water supply. *Water Science & Technology* 47(3):229–234.

United Nations (2015) The right to water. Technical report, Office of the United Nations High Commissioner for Human Rights.

van Donselaar K, van Woensel T, Broekmeulen R, Fransoo J (2006) Inventory control of perishables in supermarkets. *International Journal of Production Economics* 104(2):462–472.

WaterAid (2016) Water: At what cost? the state of the world’s water 2016. Technical report, WaterAid.

WHO, UNICEF (2017) Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines. Technical report, WHO/UNICEF Joint Monitoring Programme for Water Supply, Sanitation and Hygiene.

World Bank (2010) The Karnataka urban water sector improvement project: 24 x 7 water supply is achievable. Technical report, Water and Sanitation Program.

Zérah MH (1998) How to assess the quality dimension of urban infrastructure: The case of water supply in Delhi. *Cities* 15(4):285–290.

Zidek RA, Kolmanovsky IV (2016) Stochastic drift counteraction optimal control and enhancing convergence of value iteration. *2016 IEEE 55th Conference on Decision and Control (CDC)*, 1119–1124.