Coordinating Pricing and Empty Container Repositioning in Two-Depot Shipping Systems

Tao Lu[1], Chung-Yee Lee [*2], and Loo-Hay Lee[3]

[1]Rotterdam School of Management, Erasmus University, The Netherlands

[2]Department of Industrial Engineering and Decision Analytics, The Hong Kong University of Science and Technology, Hong Kong

[3]Department of Industrial and Systems Engineering, National University of Singapore, Singapore

December 3, 2019

Abstract

This paper studies joint decisions on pricing and empty container repositioning in two-depot shipping services with stochastic shipping demand. We formulate the problem as a stochastic dynamic programming (DP) model. The exact DP may have a high-dimensional state space due to in-transit containers. To cope with the curse of dimensionality, we develop an approximate model where the number of in-transit containers on each vessel is approximated with a fixed container flow predetermined by solving a static version of the problem. Moreover, we show that the approximate value function is $L^\natural$-concave, thereby characterizing the structure of the optimal control policy for the approximate model. With the upper bound obtained by solving the information relaxation-based dual of the exact DP, we numerically show that the control policies generated from our approximate model are close to optimal when transit times span multiple periods.

Key words: Empty container repositioning; dynamic pricing; Markov decision process; $L^\natural$-concavity; approximate dynamic programming; duality.

1 Introduction

Transportation services usually feature demand imbalance in opposite directions, which inevitably leads to unbalanced allocations of empty equipment in different locations. In ocean

---

[*]Corresponding author. Email: cylee@ust.hk

container transport, the trade imbalance has been worsening in recent decades. Based on the data from 2007 to 2012, Figure 1 shows how severe the trade imbalance is in the Europe-Asia and transpacific shipping routes, two major connections for global supply chains. In order to meet demand with sufficient empty containers in each service direction, ocean liners must redistribute their capacities by repositioning empty containers: Besides laden containers, empty containers must be moved from surplus areas to deficit areas. According to Fuller (2006), out of every 100 containers shipped from Asia to North America, 60 were sent back empty; on Asia-Europe routes, 41% went back to Asia empty. Furthermore, trade imbalance also has a significant impact on freight rates. De Oliveira (2014) reports that trade imbalance is an important factor driving the different freight rates for inward and outward journeys in a given itinerary. It is hence necessary to develop an integrated framework incorporating both repositioning and pricing decisions, in order to analyze their underlying interactions.
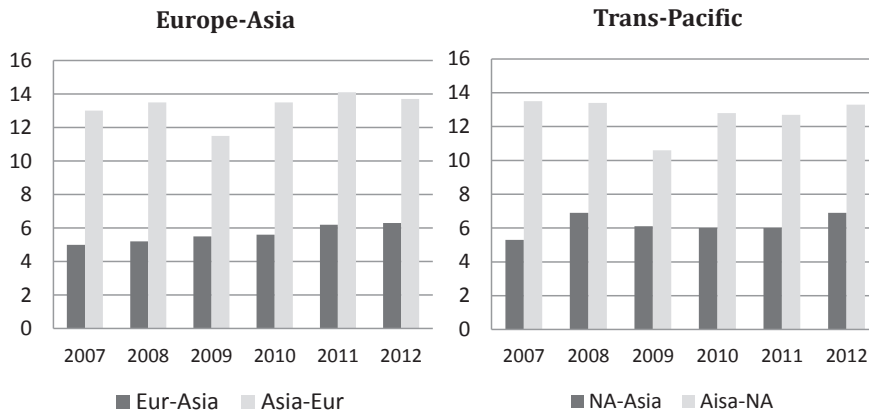

Figure 1: Containerized trade demands (in million TEUs) on two major shipping routes from 2007 to 2011 (Source: Song and Dong, 2015)

In this paper, we develop a stochastic dynamic programming (DP) model for two-depot shipping systems in which head-haul and back-haul shipping demands are random and endogenously affected by freight rates. The control variables include repositioning quantities and the pricing decisions for the voyages in both directions. In line with the literature on dynamic empty container management (e.g., Song, 2007; Ng et al., 2012), we focus on shipping routes consisting of two ports. According to Song (2007), among 1521 regular shipping services recorded by Containerization International Online, 253 are two-port shuttle services. Moreover, the two-port services can be considered as a macro-level approximation of intercontinental shipping services. For instance, the ocean liners on trans-Pacific shipping lanes are mostly concerned about the trade imbalance between two major geographic regions, i.e., Asia and North America. Viewing each port as a region, one can still apply our model to manage trans-Pacific routes on an

aggregate level.

When the transit time spans multiple periods, the exact DP model has to track the in-transit containers on every vessel, leading to a high-dimensional state space. Thus, the exact model is generally intractable due to the curse of dimensionality. To circumvent this difficulty, we develop an approximate formulation which requires only three state variables, regardless of how long the transit time is. The idea is to use some fixed number of in-transit containers to approximate the value function. The fixed number can be predetermined as the optimal container flow in a deterministic and static version of the problem.

Inspired by the recent applications of $L^\natural$-convexity/concavity in the inventory literature (e.g., Zipkin, 2008), we prove that the approximate value function is $L^\natural$-concave in a transformed state space. The $L^\natural$-concavity implies the monotonicity of the optimal solution in some of the model parameters, which enables us to characterize the interdependence between pricing and repositioning decisions. We show that the optimal prices for the approximate model are monotone in the inventory position which is defined as the number of containers at a port plus those in transit to this port. The monotone properties not only can provide general guidance for coordinating pricing with empty container management, but also can be used to reduce the search space for the optimal policies. In addition, we derive the structure of the optimal policies for the approximate model, which gives guidelines for the match-back policies adopted in practice (c.f. Lam et al., 2007). In particular, we find that it is not always optimal to maintain the flow conservation, i.e., to equate the container inflow and outflow at a port by repositioning empty containers.

To quantity the performance of our approximation, we construct an upper bound of the exact DP with the information relaxation-based duality technique (Brown et al., 2010). With this computable upper bound, we demonstrate that our approximation can generate close-to-optimal solutions and the average optimality gap is less than 2% in a variety of instances. In addition, we numerically show that the value of coordinating pricing with empty container management increases as demand imbalance escalates.

The contributions of this paper are summarized as follows: (1) To the best of our knowledge, this is the first paper that studies joint pricing and empty container repositioning decisions in a stochastic and dynamic environment. (2) We develop a novel approximation approach to overcome the curse of dimensionality arising from in-transit containers. (3) The structure of the approximate optimal policies is analytically characterized. (4) From a methodological perspective, we provide new applications of $L^\natural$-concavity and the information relaxation-based duality in the domain of maritime transportation.

The remainder of the paper is organized as follows. Section 2 reviews the literature. Section 3 describes the exact DP formulation. Section 4 presents the approximate model and analytical results. Section 5 introduces the upper bound of the exact model and Section 6 reports numerical results. Section 7 discusses several extensions and Section 8 concludes.

2   Literature Review

Empty container repositioning has long been studied in the transportation literature separately from pricing decisions. Crainic et al. (1993) propose time-space network models for the empty container allocation problem in an inland transportation system consisting of seaports, inland storage locations and customer sites. Cheung and Chen (1998) consider an ocean transportation network with demand uncertainty and develop a two-stage stochastic programming. Erera et al. (2009) adopt robust optimization techniques to address the problem in a two-stage planning framework. In addition to uncertain demand and supply, Long et al. (2012) further take into account the uncertainty in vessels' weight and space capacity and solve the problem using a two-stage stochastic programming. Although the above papers consider more complicated networks (with more than two ports) than ours, their two-stage stochastic programming frameworks assume that all uncertainties are resolved in the second stage. In reality, however, the management of empty containers is a dynamic process where demand uncertainties are sequentially resolved.

Some authors have studied dynamic empty container management without pricing using stochastic dynamic programming. Li et al. (2004) consider a single port and characterize the optimal repositioning policy, based on which Li et al. (2007) further develop a heuristic for multi-port systems. Lam et al. (2007) propose a dynamic programming formulation that minimizes the long-run average cost. More closely related to our work is the seminal paper by Song (2007) in which the author models two-port shipping systems based on a periodic-review inventory control framework. In the case of container shortfall, additional containers are leased for emergency. Ng et al. (2012) study a similar model but unsatisfied demand is backlogged. In these papers, however, shipping demands are assumed exogenous and hence pricing decisions are not addressed. Moreover, the models in Lam et al. (2007), Song (2007) and Ng et al. (2012) have implicitly assumed that the transit time between two ports is much shorter than one decision period so that their dynamic programming models have only a single state variable. In this paper, we fully relax this assumption and allow for multi-period transit times. This, however, leads to a high-dimensional stochastic DP which is generally intractable. A novel approximation approach is then proposed to reduce the state dimension and provide close-to-optimal solutions.

It is worth noting that several studies have attempted to address pricing decisions with stylized

deterministic models. Zhou and Lee (2009) study a Bertrand competition between two ocean liners operating two-port services. Empty containers must be repositioned in order to offset the imbalance of demands in the two directions. Recently, Chen et al. (2016) extend the framework of Zhou and Lee (2009) to incorporate waste shipments via empty repositioning. However, both models are static and neither one takes into account demand uncertainty.

Our work is also related to the literature on vehicle repositioning in fleet management. Gorman (2001) and King and Topaloglu (2007) assume that the number of loads on a traffic lane is a deterministic function of price and the decision maker jointly determines the price charged for each traffic lane and the number of vehicles to be relocated within the network. Topaloglu and Powell (2007) further capture demand uncertainty and model the joint optimization problem as a stochastic dynamic pricing problem. Our problem is different from the models in the fleet management literature in that our objective function includes container-based operating costs (i.e., storage and leasing costs) which are nonlinear and mirror the overage and underage costs in inventory management.

Our paper can also be positioned in the literature on inventory management with pricing. Thowsen (1975), Federgruen and Heching (1999) and Chen and Simchi-Levi (2004) are the representative works along this line. We refer interested readers to the comprehensive survey articles such as Elmaghraby and Keskinocak (2003), Chen and Simchi-Levi (2012) and Chen and Chen (2015). Because of several salient features of empty containers manage, our model departs from ordinary inventory models in several ways. First, after satisfying demand, the stock of traditional commodities is consumed, whereas empty containers are still available (at the other location). Second, in the liner service, the pricing decision for one voyage affects not only the inventory level at the origin port but also that at the destination port. Third, instead of periodically replenishing inventory through an outside source, in our problem the ocean liner determines how to redistribute empty containers within the system.

3 The Model

We consider an ocean liner providing transportation services between two ports in a finite planning horizon divided into T periods. The transit time between the two ports is L periods where $L \geq 1$ is a positive integer. The liner maintains a one-period service frequency. Since the voyages in both directions are operated in each period, the liner must deploy $2L$ vessels on the cyclic service route to maintain the service frequency. For example, when one period is equal to one week, the shipping service is operated on a weekly basis. If the transit time is one week, i.e., $L = 1$, two vessels must be deployed on the service route such that there is a vessel departing

from each port once a week.

The sequence of events is as follows: (1) Prior to the voyages in a period, the liner announces the prices for both directions; (2) demands are realized and the voyages that commenced L periods ago arrive at the destination ports; (3) based on realized demands, the liner decides how many empty containers to be repositioned and then launches new voyages in both directions.

For the t-th voyage ($t = 1, 2, ..., T+1$), let d_t^i = random demand from port i to port j and p_t^i = price charged for the voyage from port i to port j. Throughout the paper, we use indexes $i, j \in \{1, 2\}$ and $i \neq j$ to indicate the two different ports.

Like Federgruen and Heching (1999) and Chen and Simchi-Levi (2004), we assume p_t^i is selected from a finite interval $[\underline{p}_t^i, \bar{p}_t^i]$ where $\underline{p}_t^i$ (resp. $\bar{p}_t^i$) is the lowest (resp. highest) feasible price to be charged.

The expected shipping demand from port i to port j in every period t is a function of p_t^i, denoted by $D_t^i(p_t^i)$. The actual demand d_t^i is assumed to be $D_t^i(p_t^i)$ plus an additive random noise ϵ_t^i:

$$d_t^i = D_t^i(p_t^i) + \epsilon_t^i, \tag{1}$$

where the ϵ_t^i's are continuous random variables with known distributions and are independent across periods. Without loss of generality, we assume $\mathbb{E}[\epsilon_t^i] = 0$ for all i and t. In addition, we have assumed that the integrality constraints on demands and shipments are negligible. This is a reasonable assumption when the shipping line manages a large number of containers to meet substantial demand volumes.

Assumption 1. *For all $p \in [\underline{p}_t^i, \bar{p}_t^i]$, t and i, $d_t^i = D_t^i(p_t^i) + \epsilon_t^i$ is nonnegative, and $D_t^i(p_t^i)$ is finite and strictly decreasing in p_t^i.*[1]

Let λ_t^i be the expected demand from port i to port j in period t and $A_t^i(\lambda_t^i)$ denote the inverse demand function, i.e., the inverse function of $D_t^i(p_t^i)$. The expected gross revenue can therefore be written as $r_t^i(\lambda_t^i) = \lambda_t^i A_t^i(\lambda_t^i)$. Equivalent to determining p_t^i within $[\underline{p}_t^i, \bar{p}_t^i]$, we can choose λ_t^i from a given interval $[\underline{\lambda}_t^i, \bar{\lambda}_t^i]$. We make the following assumption: The expected revenue is concave in the expected demand.[2]

Assumption 2. *For all t and i, $r_t^i(\lambda)$ is concave and differentiable in λ for $\lambda \in [\underline{\lambda}_t^i, \bar{\lambda}_t^i]$.*

Note that the demands are realized after the prices are determined but before the number of empty containers to be repositioned is decided. For the analysis, it is convenient to have

---

[1] In the ocean shipping industry, the overall demand generally exhibits a low elasticity. However, individual carriers can still influence their demand by adjusting the freight rate, especially when there are substitutable services on the same route.

[2] Assumption 2 is satisfied by many commonly used demand functions, e.g., linear demand $D(p) = a - kp$, logit demand $D(p) = \frac{e^{a-kp}}{1+e^{a-kp}}$, exponential demand $D(p) = ae^{-kp}$ and log demand $D(p) = \log(a - kp)$, where $a, k > 0$.
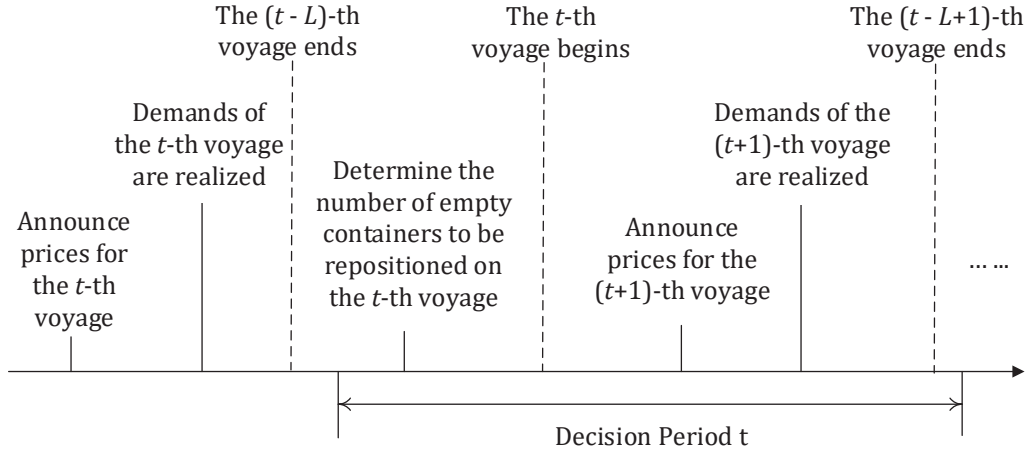
Figure 2: The sequence of events

the random noises realized at the end of each period. To achieve this, we consider events in *(decision) period* t as follows: (1) At the beginning of period t, the liner decides the number of empty containers to be repositioned based on the realized demands for the t-th voyage; (2) the liner announces p_{t+1}^i; (3) demands d_{t+1}^i's are realized. Figure 2 illustrates this sequence of events, where the starting and completion times of each voyage are indicated by dashed lines, since they are not essential to our analysis. For example, the $(t - L + 1)$-th voyage may also end prior to the announcement of p_{t+1}^i's, but all of our results would remain the same.

Laden containers are unloaded immediately upon arrival. Thus, at the beginning of period t, both laden and empty containers on the most recently completed voyage are available for the next voyage. Vessel space is assumed to be sufficient, as ocean liners are usually more concerned about the number of empty containers as the main capacity constraint.

Before a voyage begins, if the realized demand exceeds the volume that can be shipped with the liner's own empty container available at a port, additional containers will be leased[3] immediately from outside vendors to meet the demand. We assume that the liner can always lease enough containers to satisfy demands on time.[4] The leasing cost is proportional to the duration of lease. In addition, we also adopt the common assumption in the literature: All containers are functionally identical, so the liner may return any idle containers to the vendor (see, for example, Cheung and Chen, 1998; Song, 2007). This assumption implies that in a location where out-of-system containers are leased, once some containers become idle in that location, they will be automatically returned to the vendor to shorten the lease duration.[5] If

---

[3]To avoid confusion, throughout the paper, we refer to these *short-term* leased containers simply as leased containers. Ocean liners may also have long-term leased containers, but here we simply treat them as the liner's own.

[4]As mentioned by Cheung and Chen (1998), in reality, ocean liners are seldom unable to find enough containers from external sources. Therefore, they rarely reject or backlog customer orders.

[5]In practice, it is prohibitively costly to track every single container due to the huge number of containers being handled. Tracking and returning exactly every single container that was leased is therefore impossible.

there are idle containers after the voyage has begun, an inventory holding cost will be incurred for each idle container per period. The holding cost refers to the expenses incurred for storing idle containers in the port terminal/inland container yard.[6]

Accordingly, define the following cost parameters for each period t:

b_t^i = leasing cost per period for one unit of container at port i;

h_t^i = inventory holding cost per period for one unit of container at port i;

c_t^f = one-time cost for handling one unit of laden container on the t-th voyage;

c_t^e = one-time cost for handling one unit of empty container on the t-th voyage.

Assumption 3. $c_t^e > h_t^i$ *for all t and i.*

The above assumption requires that it be cheaper to hold containers in inland depots than to load them on board. If this assumption fails, the liner would purposely load empty containers on board to reduce inventory holding costs incurred in inland depots. This is clearly not common in reality, as the cost for handling empty containers on a voyage is normally higher than that for storing them inland. Throughout the paper, we will assume that Assumptions 1-3 are satisfied unless otherwise specified. Let

u_t^i = number of empty containers to be repositioned from port i to port j on the t-th voyage;

$z_{0,t}^i$ = number of available empty containers owned by the liner at the beginning of period t at port i (note that when location i has a deficit capacity, z_t^i will take a negative value, the absolute value of which represents the number of containers being leased from location i);

$z_{l,t}^i$ = number of (both laden and empty) containers in transit that will arrive at port i in period $t + l$, where $l = 1, 2, ..., L$.

By definition, it follows that

$$z_{L,t}^i = d_t^j + u_t^j \text{ for } i, j \in \{1, 2\} \text{ and } i \neq j \tag{2}$$

which represents the total number of containers dispatched from port j in period t, including laden containers (i.e., shipment demand d_t^j) and empty ones (i.e., u_t^j).

The system dynamics are characterized by $2L + 2$ state variables, i.e., inventory levels $\mathbf{z}_{l,t} = (z_{l,t}^1, z_{l,t}^2)$ where $l = 0, 1, ..., L - 1$ and realized demands $\mathbf{d}_t = (d_t^1, d_t^2)$, together with the following

---

Even though containers come in different types, e.g., different sizes, given a sufficiently large volume (which we have implicitly assumed by using continuous variables to count containers), it is reasonable to assume that the liner can return containers of a particular size once there are idle containers on hand.

[6]The liner may also face a problem of whether to keep the idle containers in the port terminal or move them to the inland container yard. This depends on the terminal operator's pricing scheme for container storage, the inland transportation cost, etc. See Lee and Yu (2012) for a study pertaining to this issue. In this paper, however, we do not consider inland container flows and refer to the cost incurred for storing idle containers inland as the inventory holding cost.

equations:

$$z_{0,t+1}^i = z_{0,t}^i - z_{L,t}^j + z_{1,t}^i \quad , i,j \in \{1,2\} \text{ and } i \neq j \tag{3}$$

$$z_{l,t+1}^i = z_{l+1,t}^i, \quad \text{for } i = 1,2, \text{ and } l = 1,2,...,L-1 \tag{4}$$

$$d_{t+1}^i = \lambda_{t+1}^i + \epsilon_{t+1}^i \text{ for } i = 1,2. \tag{5}$$

In addition, we assume that the total number of containers owned by the liner, denoted by N, is fixed during the planning horizon. Hence, $\sum_{i=1}^2 \sum_{l=0}^{L-1} z_{l,t}^i = N$ for all t. Consequently, it is sufficient to use $2L+1$ state variables to describe the system dynamics. For ease of exposition, we will continue presenting our model with $2L+2$ state variables but use $2L+1$ state variables in the numerical study.
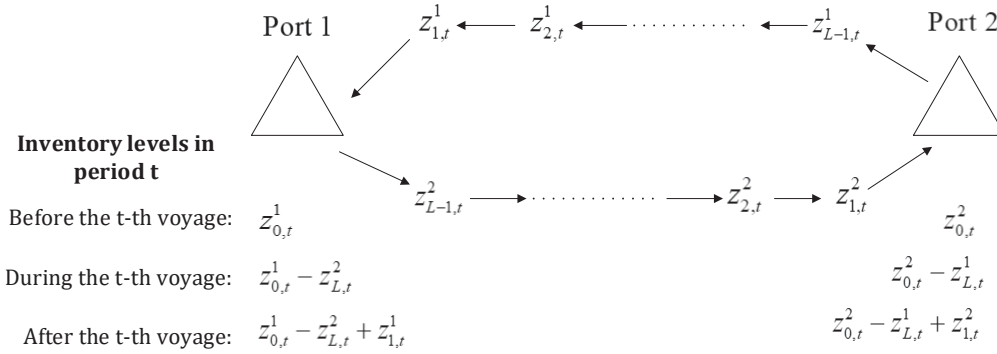

Figure 3: Dynamics of inventory levels

Figure 3 illustrates how the inventory levels in the two locations evolve over time. We allow inventory levels to be negative to capture the deficit scenario. That is, $z_{0,t}^i - u_t^i - d_t^i < 0$ indicates that there are containers being leased at port i. For example, consider $N = 30, L = 1$ and at the beginning of period t, $z_{0,t}^1 = 10$ at port 1 and hence the number of containers at port 2 is given by $z_{0,t}^2 = N - z_{0,t}^1 = 20$ since there are no containers in transit when $L = 1$. During the t-th voyages, suppose that we ship 15 units of containers in each direction, i.e., $z_{L,t}^1 = z_{L,t}^2 = 15$. During this voyage, in addition to the 30 units of containers at sea, the inventory level at port 1 equals $z_{0,t}^1 - z_{L,t}^2 = -5$, indicating 5 units of containers being leased, and the inventory level at port 2 is 5. At the end of period t, we will have $z_{0,t+1}^1 = -5 + 15 = 10$ at port 1 and $z_{0,t+1}^2 = N - z_{0,t+1}^1 = 20$ at port 2, as the leased containers at port 1 have been returned once extra containers become idle.

Let $G_t^i(x) = h_t^i \cdot (x)^+ + b_t^i \cdot (x)^-$, where $(x)^+ = \max\{x, 0\}$ and $(x)^- = \max\{-x, 0\}$. The leasing and inventory holding costs in period t at port i is then given by $G_t^i(z_{0,t}^i - z_{L,t}^j)$, since $z_{0,t}^i - z_{L,t}^j$ represents the inventory level at port i during the t-th period.

9

For the timing of cash flow, for simplicity, we assume that the revenue and container handling costs of a voyage are respectively received and paid at the end of the voyage. The container leasing and holding costs in period t are incurred once the t-th voyage begins. Let $0 < \alpha \leq 1$ be the discount factor. The liner's objective is to maximize the expected total profit over the entire planning horizon.

We let the expected demands $\boldsymbol{\lambda}_{t+1} = (\lambda_{t+1}^1, \lambda_{t+1}^2)$ and the number of containers loaded on the t-th voyages $\mathbf{z}_{L,t} = (z_{L,t}^1, z_{L,t}^2)$ be the decision variables in period t, where $\lambda_{t+1}^i \in [\underline{\lambda}_{t+1}^i, \bar{\lambda}_{t+1}^i]$ and $z_{L,t}^j \geq d_t^i$ since $u_t^i \geq 0$. Define $R_{t+1}(\boldsymbol{\lambda}_{t+1}) = \sum_{i=1}^2 [r_{t+1}^i(\lambda_{t+1}^i) - c_{t+1}^f \lambda_{t+1}^i]$ as the net revenue from meeting demands on the $(t+1)$-th voyage. We use $J_t(\mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t)$ to denote the profit-to-go function for period t. For $t = 1, 2, ..., T$, the DP recursion can then be written as

$$J_t(\mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t) = \max_{\boldsymbol{\lambda}_{t+1}, \mathbf{z}_{L,t}} \quad f_t(\mathbf{z}_{L,t}, \boldsymbol{\lambda}_{t+1}, \mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t) + c_t^e \cdot (d_t^1 + d_t^2)$$
$$s.t. \quad \lambda_{t+1}^i \in [\underline{\lambda}_{t+1}^i, \bar{\lambda}_{t+1}^i], \quad z_{L,t}^j \geq d_t^i, \text{ for } i = 1, 2, \tag{6}$$

where

$$f_t(\mathbf{z}_{L,t}, \boldsymbol{\lambda}_{t+1}, \mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t) = \alpha R_{t+1}(\boldsymbol{\lambda}_{t+1}) - \sum_{i=1}^2 \left(c_t^e z_{L,t}^i + G_t^i(z_{0,t}^i - z_{L,t}^j) \right)$$
$$+ \alpha \mathbb{E} J_{t+1}(\mathbf{z}_{0,t+1}, ..., \mathbf{z}_{L-1,t+1}, \mathbf{d}_{t+1}). \tag{7}$$

The state variables in period $t + 1$, i.e., $(\mathbf{z}_{0,t+1}, \mathbf{z}_{1,t+1}, ..., \mathbf{z}_{L-1,t+1}, \mathbf{d}_{t+1})$, are determined by decision variables $\boldsymbol{\lambda}_{t+1}, \mathbf{z}_{L,t}$, random noises $\boldsymbol{\epsilon}_t$ and the state variables in period t according to equations (3), (4) and (5).

In the expression, the net revenue from the $(t + 1)$-th voyage is counted in period t, which is given by $\alpha R_{t+1}(\boldsymbol{\lambda}_{t+1})$. The repositioning cost on the t-th voyage is given by $c_t^e \sum_{i=1}^2 u_t^i = c_t^e \left(\sum_{i=1}^2 z_{L,t}^i - \sum_{i=1}^2 d_t^i \right)$ where the term $c_t^e \sum_{i=1}^2 d_t^i$ is removed from the reward function f_t to be optimized. As a termination condition, we set $J_{T+1}(\mathbf{z}_{0,T+1}, \mathbf{z}_{1,T+1}, ..., \mathbf{z}_{L-1,T+1}, \mathbf{d}_{T+1}) = -\sum_{i=1}^2 G_{T+1}^i(z_{0,T+1}^i - d_{T+1}^i)$ so that the container holding and leasing costs for the $(T + 1)$-th voyages are included but no more voyages start after the $(T + 1)$-th voyages.

We close the subsection by remarking that Song (2007) has also studied a DP model for two-depot shipping systems like ours. Our model is more general than his in two important aspects. First, in Song's model demand is exogenous and the liner only determines repositioning quantity, whereas we endogenize the demand by incorporating pricing decisions. Second, unlike our model where the inventory cost and leasing cost are charged according to the inventory levels *during* the voyage, Song counts the costs based on the end-of-voyage inventory positions in each period. In his model, a single state variable is enough to describe the dynamics, but this formulation is only

suitable for the case where the transit time is very short relative to one period (i.e., $L << 1$). Relaxing this assumption requires a larger state space with $2L + 1$ state variables. To cope with the curse of dimensionality, in the next section we will propose an approximate formulation in which the dimension of the state space can be reduced to three, regardless of the value of L.

4 The Approximate Model

4.1 State Dimension Reduction

Our approximation method aims to reduce the dimension of the state space, which is a major difficulty in solving the exact DP (6). Following the often used transformation in the inventory management literature (e.g., Porteus, 2002), we can use the inventory position (i.e., on-hand inventory level plus orders in transit) to replace the inventory level at each port. In our system, the inventory position at port i can be defined as $x_t^i = \sum_{l=0}^{L-1} z_{l,t}^i$, i.e., the number of containers at port i plus the containers in transit to port i. However, this does *not* resolve the curse of dimensionality, as we must still keep track of the in-transit containers on each vessel $z_{l,t}^i$ during the recursion to calculate the container leasing and holding costs $G_t^i(x_t^i - \sum_{l=1}^{L-1} z_{l,t}^i - z_{L,t}^j)$. Moreover, the number of in-transit containers on each vessel can be dynamically controlled through pricing and repositioning decisions.

Note that $z_{l,t}^i = d_{t+l-L}^j + u_{t+l-L}^j$, i.e., the number of containers sent from port j to port i in period $t + l - L$, which depends on the decision variables $\boldsymbol{\lambda}_{t+l-L}$ and u_{t+l-L}^i. We can therefore approximate each $z_{l,t}^i$ with some fixed number $\bar{z}_{l,t}^i = \bar{d}_{t+l-L}^i + \bar{u}_{t+l-L}^i$, where the values of $\bar{d}_{t+l-L}^i$ and $\bar{u}_{t+l-L}^i$ can be obtained beforehand by solving some outer optimization problems with deterministic demand. In particular, we consider the following deterministic problem:

$$\max_{\lambda_t^i \in [\underline{\lambda}_t^i, \bar{\lambda}_t^i]} R_t(\lambda_t) - c_t^e \cdot |\lambda_t^1 - \lambda_t^2|, \tag{8}$$

where the first term captures the net revenue of the t-th voyage and in the second term we require that the demand imbalance be exactly offset through empty container repositioning at a unit cost of c_t^e. Problem (8) can be viewed as a static version of our original problem, which is in the same spirit as that studied in Zhou and Lee (2009). Let $\bar{\lambda}_t^1$ and $\bar{\lambda}_t^2$ denote the optimal solution to problem (8). We can then set $\bar{d}_{t+l-L}^i = \bar{\lambda}_{t+l-L}^i$ and $\bar{u}_{t+l-L}^i = (\bar{\lambda}_{t+1-L}^j - \bar{\lambda}_{t+1-L}^i)^+$ and use $\bar{z}_{l,t}^i = \bar{d}_{t+l-L}^i + \bar{u}_{t+l-L}^i$ to approximate $z_{l,t}^i$. In other words, we use the optimal container flow derived from the static problem (8) to approximate the number of in-transit containers.[7]

---

[7]Our approximation approach is inspired by Federgruen and Heching (1999, 2002) where the authors propose the idea of using a fixed price path derived from a deterministic version of the problem to approximate in-transit inventories.

We define an approximate cost function

$$\hat{G}_t^i(x_t^i - z_{L,t}^j) = G_t^i(x_t^i - \sum_{l=1}^{L-1} \bar{z}_{l,t}^i - z_{L,t}^j)$$

and use it to approximate the value function in the exact DP (6). Since $x_t^1 + x_t^2 = N$ for all t, we will hereafter simply use x_t to denote the inventory position at port 1 and the inventory position at port 2 is then given by $N - x_t$. For $t = 1, 2, ..., T$, using the relation $z_{L,t}^i = d_t^j + u_t^j$, our approximate DP recursion can be written as

$$
\begin{aligned}
J_t^A(x_t, \mathbf{d}_t) &= \max_{\lambda_{t+1}^i \in [\underline{\lambda}_{t+1}^i, \bar{\lambda}_{t+1}^i], u_t^i \geq 0} f_t^A(\mathbf{u}_t, \boldsymbol{\lambda}_{t+1}, x_t, \mathbf{d}_t) \\
&= \max_{\lambda_{t+1}^i \in [\underline{\lambda}_{t+1}^i, \bar{\lambda}_{t+1}^i], u_t^i \geq 0} \{\alpha R_{t+1}(\boldsymbol{\lambda}_{t+1}) - c_t^e \cdot (u_t^1 + u_t^2) \\
&\quad - \hat{G}_t^1(x_t - u_t^1 - d_t^1) - \hat{G}_t^2(N - x_t - u_t^2 - d_t^2) + \alpha \mathbb{E} J_{t+1}^A(x_{t+1}, \mathbf{d}_{t+1})\},
\end{aligned}
\tag{9}
$$

where

$$x_{t+1} = x_t - u_t^1 - d_t^1 + u_t^2 + d_t^2, \tag{10}$$

and $\mathbf{d}_{t+1}$ is determined by (5). The termination condition is then rewritten as $J_{T+1}^A(x_{T+1}, \mathbf{d}_{T+1}) = -\hat{G}_{T+1}^1(x_{T+1} - d_{T+1}^1) - \hat{G}_{T+1}^2(N - x_{T+1} - d_{T+1}^2)$.

Note that the approximate formulation involves three state variables x_t, d_t^1 and d_t^2. The inventory level x_t alone is not sufficient because repositioning quantities are determined *after* the actual demands are received. That is, the repositioning decision is contingent on d_t^1 and d_t^2 as well. Intuitively, the ocean liner would reposition fewer (resp. more) containers if the realized demand in the same direction turns out to be higher (resp. lower).

We can use the optimal control policy for problem (9) as an approximate solution to the exact DP (6). The merit of our approximate model is that no matter how long the transit time is, the proposed approximation has only three state variables whereas the exact model needs $2L + 1$ state variables with L-period transit times, with the understanding that one more state variable increases the state space in an exponential manner! Moreover, if the transit time is one period, i.e., $L = 1$, the approximate model is equivalent to the exact formulation (6), since we only approximate in-transit containers which appear in the value function only when $L > 1$.

4.2 Analysis of the Approximate Model

Although we have reduced the state space to three dimensions in the approximate model (9), it remains challenging to analyze the structure of the optimal policy to this model.

4.2.1 Preliminaries

To derive the monotonicity of optimal policies, we apply the concept of $L^\natural$-concavity (e.g., Pang et al., 2012). Interested readers are referred to Appendix A for formal statements of its definition and properties. $L^\natural$-concavity implies ordinary concavity and supermodularity, thus allowing us to characterize how the optimal decision is monotonic in multi-dimensional parameters. For example, if we maximize function $g(\mathbf{v}, \zeta)$ over $\zeta \geq 0$ where $\mathbf{v}$ is a vector consisting of multiple parameters, roughly speaking, the $L^\natural$-concavity of $g(\mathbf{v}, \zeta)$ implies that the optimal solution $\zeta(\mathbf{v})$ is nondecreasing in $\mathbf{v}$. In order to obtain the $L^\natural$-concavity of the value function, we transform the original state variables as follows. Define

$$
\mathbf{v}_t = \begin{bmatrix} 0 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ d_t^1 \\ d_t^2 \end{bmatrix} = \begin{bmatrix} -d_t^1 \\ x_t - d_t^1 \\ x_t - d_t^1 + d_t^2 \end{bmatrix}
$$

as the new state vector. Note that the state space $\mathcal{V} = \{\mathbf{v} : v^1 \leq 0, v^2 \leq v^3\}$ forms a lattice, as the inequality involving more than one variable has exactly two variables with opposite signs (see Example 2.2.7 in Topkis (1998)). Although this transformation is performed mainly for technical reasons, the state vector $\mathbf{v}$ does have some physical meanings: v_t^2 represents the inventory position at port 1 deducting the number of containers that have been reserved for the t-th voyage, and v_t^3 indicates the net inventory position at port 1 after the inbound and outbound containers reserved for the t-th voyage are taken into account.

Then, define

$$
\begin{aligned}
y_t^1 &= v_t^3 - u_t^1 + u_t^2, \\
y_t^2 &= y_t^1 + u_t^1, \\
y_t^3 &= y_t^1 + \lambda_{t+1}^2.
\end{aligned}
$$

Note that y_t^1 is a critical variable in our problem, and $y_t^1 = x_{t+1}$, i.e., it equals the inventory position at port 1 at the end of the t-th voyage/at the beginning of the $(t+1)$-th voyage. We will refer to y_t^1 as the end-of-voyage inventory position (at port 1). Note that the inventory position at port 2 is simply given by $N - y_t^1$.

Accordingly, the approximate DP formulation (9) can be rewritten as

$$
\begin{aligned}
J_t^A(\mathbf{v}_t) = \max_{(u_t^1, y_t^1, y_t^2, y_t^3, \lambda_{t+1}^1) \in \mathcal{A}} &\{ \alpha R_{t+1}(\lambda_{t+1}^1, y_t^3 - y_t^1) - c_t^e \cdot (u_t^1 + y_t^2) - \hat{G}_t^1(v_t^2 - u_t^1) \\
&- \hat{G}_t^2(N - y_t^2 + v_t^1) + \alpha \mathbb{E} J_{t+1}^A(\mathbf{v}_{t+1}) \} + c_t^e v_t^3, \text{ for } t = 1, 2, ..., T
\end{aligned} \tag{11}
$$

13

and, for the last period,

$$J_{T+1}^A(\mathbf{v}_{T+1}) = -\hat{G}_{T+1}^1(v_{T+1}^2) - \hat{G}_{T+1}^2(N + v_{T+1}^1 - v_{T+1}^3), \quad (12)$$

where the system dynamics translates to

$$\mathbf{v}_{t+1} = \begin{bmatrix} -\lambda_{t+1}^1 - \epsilon_{t+1}^1 \\ v_t^3 - u_t^1 + u_t^2 - \lambda_{t+1}^1 - \epsilon_{t+1}^1 \\ v_t^3 - u_t^1 + u_t^2 - \lambda_{t+1}^1 - \epsilon_{t+1}^1 + \lambda_{t+1}^2 + \epsilon_{t+1}^2 \end{bmatrix}$$
$$= (0, y_t^1, y_t^3 + \epsilon_{t+1}^2)^T - (\lambda_{t+1}^1 + \epsilon_{t+1}^1)\mathbf{e}.$$

The action space $\mathcal{A} = \{(u_t^1, y_t^1, y_t^2, y_t^3, \lambda_{t+1}^1) : u_t^1 \geq 0, y_t^2 \geq v_t^3, u_t^1 + y_t^1 = y_t^2, y_t^3 - y_t^1 \in [\underline{\lambda}_{t+1}^2, \bar{\lambda}_{t+1}^2], \lambda_{t+1}^1 \in [\underline{\lambda}_{t+1}^1, \bar{\lambda}_{t+1}^1]\}$. $\mathcal{A}$ is *nonlattice* due to the constraint $u_t^1 + y_t^1 = y_t^2$. The nonlattice structure gives rise to another analytical difficulty, because a generic way to show the preservation of $L^\natural$-concavity (like supermodularity) under maximization requires the constraint set to be lattice. [8]

In this paper, we circumvent the nonlattice structure by dividing the decisions into two stages: The liner determines firstly the repositioning quantity corresponding to variables u_t^1, y_t^1 and y_t^2, then the prices for the next voyage corresponding to variables λ_{t+1}^1 and y_t^3. In the second-stage decision, for any given (u_t^1, y_t^1, y_t^2), we find that the maximization over λ_{t+1}^1 and y_t^3 depends on other parameters only through y_t^1. In other words, the pricing decision is made based on the inventory position for the upcoming voyage, given any repositioning quantities. Define

$$H_t(y_t^1) = \max_{(y_t^3, \lambda_{t+1}^1) \in A(y_t^1)} \alpha\{R_{t+1}(\lambda_{t+1}^1, y_t^3 - y_t^1) + \mathbb{E}J_{t+1}^A[(0, y_t^1, y_t^3 + \epsilon_{t+1}^2)^T - (\lambda_{t+1}^1 + \epsilon_{t+1}^1)\mathbf{e}]\}, \quad (13)$$

where $A(y_t^1) = \{(y_t^3, \lambda_{t+1}^1) : y_t^3 - y_t^1 \in [\underline{\lambda}_{t+1}^2, \bar{\lambda}_{t+1}^2], \lambda_{t+1}^1 \in [\underline{\lambda}_{t+1}^1, \bar{\lambda}_{t+1}^1]\}$. The function H_t serves as a key connection between pricing and repositioning decisions.

In the first stage, we solve

$$J_t^A(\mathbf{v}_t) = \max\{H_t(y_t^1) - c_t^e(u_t^1 + y_t^2) - \hat{G}_t^1(v_t^2 - u_t^1) - \hat{G}_t^2(N - y_t^2 + v_t^1)\} + c_t^e v_t^3$$

$$s.t. \quad y_t^2 = y_t^1 + u_t^1, \quad (14)$$

$$y_t^2 \geq v_t^3, u_t^1 \geq 0.$$

---

[8]Recently, Chen et al. (2013) identify some sufficient conditions under which the $L^\natural$-concavity can be preserved even when the constraint set is nonlattice. Their results require that the value function is parametrized by two-dimensional state vectors. Unfortunately, our state vector has three dimensions.

Eliminating y_t^1 with the equality constraint $y_t^2 = y_t^1 + u_t^1$, the feasible region of (14) then becomes lattice, leading to the first-stage repositioning decision.

$$
\begin{aligned}
J_t^A(\mathbf{v}_t) = \max_{y_t^2 \geq v_t^3, u_t^1 \geq 0} \{ & H_t(y_t^2 - u_t^1) - c_t^e \cdot (u_t^1 + y_t^2) - \hat{G}_t^1(v_t^2 - u_t^1) \\
& - \hat{G}_t^2(N - y_t^2 + v_t^1) \} + c_t^e v_t^3
\end{aligned}
\tag{15}
$$

With the two-stage reformulation defined above, it can be shown that the value function of our problem is indeed $L^\natural$-concave in the transformed state variables. We relegate all technical proofs to the appendices.

Lemma 1. *For $t = 1, 2, ..., T+1$, $H_t(y)$ is concave in y, and $J_t^A(\mathbf{v})$ is $L^\natural$-concave in $\mathbf{v}$.*

To establish the $L^\natural$-concavity, we have made use of the fact that the pricing decision is affected by other variables only through y_t^1. It should be noted that, in general, with a nonlattice constraint set and a three-dimensional state space, the $L^\natural$-concavity may not be preserved. Chen et al. (2013) have provided a counterexample. Fortunately, we are able to prove the $L^\natural$-concavity of J_t^A by exploiting the special structure of our problem. With the two-stage treatment above, the $L^\natural$-concavity of J_t^A in fact follows as long as J_{t+1}^A is jointly concave. From this perspective, the $L^\natural$-concavity is due to the inherent nature of our problem, rather than preservation under the DP recursion.[9]

4.2.2 Monotone Properties of the Optimal Policy

The following theorem characterizes the monotone properties of the optimal price vector with respect to the inventory position, where we denote by $\lambda_{t+1}^i(y)$ the optimal expected demand from port i to port j given that the end-of-voyage inventory position at port 1 is equal to y.

Theorem 1. *Given any repositioning quantities (u_t^1, u_t^2), the optimal price vector $(p_{t+1}^{1*}, p_{t+1}^{2*})$ depends only on $y_t^1 = v_t^3 - u_t^1 + u_t^2$, i.e., the end-of-voyage inventory position. Furthermore, for $\omega > 0$*

$$
0 \leq \lambda_{t+1}^1(y_t^1 + \omega) - \lambda_{t+1}^1(y_t^1) \leq \omega,
$$

$$
-\omega \leq \lambda_{t+1}^2(y_t^1 + \omega) - \lambda_{t+1}^2(y_t^1) \leq 0.
$$

That is, p_{t+1}^{1*} (resp. p_{t+1}^{2*}) is nonincreasing (resp. nondecreasing) in y_t^1 with bounded sensitivities.

Theorem 1 implies that the optimal pricing and repositioning quantities should be interdependent. With more empty containers available at port 1, a lower price should be charged for the voyage from port 1 to port 2 to attract more demand in that direction. Likewise, a higher price

---

[9]We would like to thank Xin Chen for pointing out this issue, which helps us clarify the implications behind the proof of $L^\natural$-concavity.

should be charged for the reverse direction to reduce the number of laden containers coming back to port 1. In addition, our results indicate that despite the complex evolution of (laden and empty) container flows in the system, to determine $\boldsymbol{p}_{t+1}$, the manager only needs to base the pricing decision on the inventory position at the end of the t-th voyage.

The following theorem states the monotone properties of the optimal repositioning quantity, where we denote by $u_t^{1*}(x_t, d_t^1, d_t^2)$ the optimal quantity repositioned from port 1 to port 2 given any state (x_t, d_t^1, d_t^2). Note that the optimal repositioning quantity from port 2 to port 1 has the same properties regarding the corresponding state, as we can simply swap the ports' indices in the model.

Theorem 2. *(i) For any $\omega > 0$, the optimal repositioning quantity u_t^{1*} satisfies*

$$
\begin{aligned}
0 \ &\le u_t^{1*}(x_t, d_t^1, d_t^2 + \omega) - u_t^{1*}(x_t, d_t^1, d_t^2) \\
&\le u_t^{1*}(x_t + \omega, d_t^1, d_t^2) - u_t^{1*}(x_t, d_t^1, d_t^2) \\
&\le u_t^{1*}(x_t, d_t^1 - \omega, d_t^2) - u_t^{1*}(x_t, d_t^1, d_t^2) \le \omega.
\end{aligned}
$$

(ii) Assuming that ω is such that $u_t^{2*} = 0$ for all states considered in the above inequalities, the third inequality in part (i) holds with equality, i.e., $u_t^{1*}(x_t + \omega, d_t^1, d_t^2) = u_t^{1*}(x_t, d_t^1 - \omega, d_t^2)$. That is, u_t^{1*} is affected by x_t and d_t^1 only through $x_t - d_t^1$.

Part (i) of Theorem 2 implies that more empty containers should be repositioned from port 1 to port 2, if we have more (resp. less) shipping demand for the voyage in the opposite (resp. same) direction or more empty containers are available at port 1 at the beginning of voyage. More interestingly, the repositioning quantity is more (resp. less) sensitive to the number of empty containers available at the origin port than to the shipping demand in the opposite (resp. same) direction. Furthermore, all of the sensitivities are bounded by one. Assuming differentiability, Theorem 2 implies $0 \le \frac{\partial u_t^{1*}}{\partial d_t^2} \le \frac{\partial u_t^{1*}}{\partial x_t} \le -\frac{\partial u_t^{1*}}{\partial d_t^1} \le 1$.

Part (ii) of Theorem 2 states the sensitivities of u_t^{1*} when the perturbation in state variables does not change the optimal repositioning direction. With a fixed d_t^2, the repositioning quantity from port 1 to port 2 will remain the same as long as the term $x_t - d_t^1$ is unchanged, assuming that the repositioning direction is always from port 1 to port 2 in the optimal solution. That is, the effect of a higher inventory position at port 1 can be offset by an increase in the demand from port 1 to port 2.

Our result is in notable contrast to that of Song (2007). In Song (2007), the optimal repositioning quantity depends only on the end-of-voyage inventory level $x_t - d_t^1 + d_t^2$. In other words, the number of empty containers to be repositioned remains unchanged if d_t^1 increases and d_t^2 decreases by the same amount. In our setting, however, we show that u_t^{1*} is more sensitive to d_t^1

than to d_t^2. The reason is that our model captures the time lag in transporting containers. Intuitively, when empty containers need to be shipped from port 1 to 2, d_t^1 affects the repositioning decision more immediately than d_t^2, as d_t^2 will not arrive at port 1 until the end of the period.

From the computational perspective, we note that Theorems 1 and 2 can be iteratively leveraged to dramatically reduce the search space for the optimal decisions. For example, to solve the DP, we need to find the optimal control for every possible state (x_t, d_t^1, d_t^2). Once we find $u_t^{1*}(x_t, d_t^1, d_t^2)$ under some state (x_t, d_t^1, d_t^2), it suffices to search for the optimal repositioning quantity under another state $(x_t + \omega, d_t^1, d_t^2)$ where $\omega > 0$ between $u_t^{1*}(x_t, d_t^1, d_t^2)$ and $u_t^{1*}(x_t, d_t^1, d_t^2) + \omega$.

4.2.3 The Structure of the Approximate Optimal Policy

As the pricing decision is affected only by the end-of-voyage inventory position y_t^1, we are interested in representing the optimal policy in terms of y_t^{1*} where $y_t^{1*} = x_t - d_t^1 + d_t^2 - u_t^{1*} + u_t^{2*}$. In what follows, we will focus on repositioning decisions, and the optimal prices are determined once the repositioning quantity is chosen. Let u_t be the *net* repositioning quantity: $u_t = u_t^1 - u_t^2$. We show in Appendix C that it is not optimal to simultaneously transport empty containers in both directions. That is, at most one of u_t^1 and u_t^2 is positive. We can hence rewrite problem (15) as an unconstrained optimization over u_t:

$$\max_{u_t} H_t(v_t^3 - u_t) - W_t(u_t, \mathbf{v}_t), \tag{16}$$

where $W_t(u_t, \mathbf{v}_t) = c_t^e \cdot |u_t| + \hat{G}_t^1(v_2 - [u_t]^+) + \hat{G}_t^2(N + v_t^1 - v_t^3 - [u_t]^-)$.

It is not difficult to verify that the cost term $W_t(u_t, \mathbf{v})$ is piecewise convex in u_t. Together with the concavity of H_t, the first-order condition guarantees the global optimality. To obtain more explicit characterizations of the optimal policy, we need a mild assumption:

Assumption 4. *Over all periods, it is not optimal to lease more empty containers than the current shortfall and transport the extra ones to the other location.*

Assumption 4 excludes the situation where the liner leases extra containers from one location and repositions them to the other location. This rarely happens in practice because (1) the liner normally has easy access to container leasing companies in most port regions of the world; and (2) doing so will incur a significant cost for handling extra empty containers during the voyage. Clearly, a sufficient condition for Assumption 4 to hold is that the cost parameters c_t^e and b_t^i are time-invariant and $b_t^1 = b_t^2$ (i.e., leasing containers in two locations are equally costly).

Under Assumption 4, the optimal net repositioning quantity u_t^* must be within the interval

$-(N - x_t - d_t^2)^+ \leq u_t^* \leq (x_t - d_t^1)^+$. We can therefore restrict our attention to the repositioning quantities that do not exceed the number of on-hand empty containers, excluding those being reserved for the upcoming voyage. In some sense, we can view $-(N - x_t - d_t^2)^+ \leq u_t \leq (x_t - d_t^1)^+$ as a state-dependent capacity constraint.

Clearly, $u_t^* = 0$ if $d_t^1 \geq x_t$ and $d_t^2 \geq N - x_t$; $u_t^* \geq 0$ if $d_t^1 < x_t$ and $d_t^2 \geq N - x_t$; $u_t^* \leq 0$ if $d_t^1 \geq x_t$ and $d_t^2 < N - x_t$. In the above three cases, the repositioning direction is simply due to the capacity constraint. For the remaining case where $d_t^1 < x_t$ and $d_t^2 < N - x_t$, by examining first-order optimality conditions of (16), we can conclude that there exist two thresholds $\bar{v}_t$ and $\underline{v}_t$ such that (i) $u_t^* = 0$ if $\underline{v}_t \leq v_t^3 \leq \bar{v}_t$; (ii) $u_t^* \geq 0$ when $v_t^3 > \bar{v}_t$; and (iii) $u_t^* \leq 0$ when $v_t^3 < \underline{v}_t$. Recall that $v_t^3 = x_t - d_t^1 + d_t^2$. The repositioning direction depends on the magnitude of the realized demand imbalance $d_t^1 - d_t^2$. Detailed mathematical discussions can be found in the proof of Theorem 3 in Appendix D.
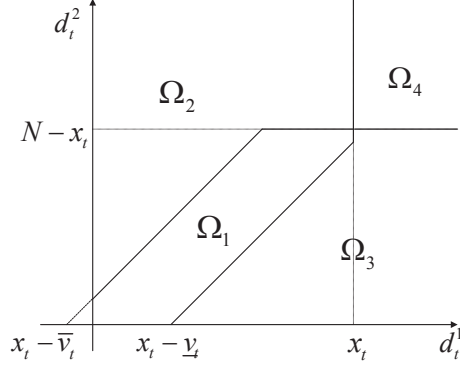

Figure 4: The state segmentation according to the repositioning direction

The overall state space, according to the sign of u_t^*, is segmented into four regions below (see also Figure 4).

$$\Omega_1 = \left\{ (x_t, d_t^1, d_t^2) : \begin{array}{l} x_t - \bar{v}_t \leq d_t^1 - d_t^2 \leq x_t - \underline{v}_t, \\ d_t^1 < x_t, d_t^2 < N - x_t \end{array} \right\}$$

$$\Omega_2 = \left\{ (x_t, d_t^1, d_t^2) : \begin{array}{l} d_t^1 - d_t^2 < x_t - \bar{v}_t, \\ d_t^1 < x_t, d_t^2 < N - x_t \end{array} \right\} \bigcup \left\{ (x_t, d_t^1, d_t^2) : d_t^1 < x_t, d_t^2 \geq N - x_t \right\}$$

$$\Omega_3 = \left\{ (x_t, d_t^1, d_t^2) : \begin{array}{l} d_t^1 - d_t^2 > x_t - \underline{v}_t, \\ d_t^1 < x_t, d_t^2 < N - x_t \end{array} \right\} \bigcup \left\{ (x_t, d_t^1, d_t^2) : d_t^1 \geq x_t, d_t^2 < N - x_t \right\}$$

$$\Omega_4 = \left\{ (x_t, d_t^1, d_t^2) : d_t^1 \geq x_t, d_t^2 \geq N - x_t \right\}$$

As three state variables are involved, for ease of exposition, we illustrate the state segmentation with a d_t^1-d_t^2 coordinate system where x_t takes a fixed value, as shown in Figure 4. In the

following theorem, we characterize the optimal repositioning quantity in each of the segments Ω_i. The pricing decision is then determined by the end-of-voyage inventory positions $x_t - d_t^1 + d_t^2 - u_t^*$.

Theorem 3. *Under Assumption 4, for any given state (x_t, d_t^1, d_t^2), the optimal policy can be characterized by two target inventory positions (s_{Ot}^*, s_{It}^*) and a price vector $\mathbf{p}_{t+1}^*(y)$, where y is the end-of-voyage inventory position. The optimal decision in period t is given by one of the following cases:*

(I) *If $(x_t, d_t^1, d_t^2) \in \Omega_1 \bigcup \Omega_4$, reposition nothing, i.e., $u_t^* = 0$ and charge $\mathbf{p}_{t+1}^*(x_t - d_t^1 + d_t^2)$.*

(II) *If $(x_t, d_t^1, d_t^2) \in \Omega_2$, the net repositioning quantity is given by*

$$
u_t^* = \begin{cases} 0 & \text{if } x_t - d_t^1 + d_t^2 \leq \bar{v}_t \\ x_t - d_t^1 + d_t^2 - s_{Ot}^* & \text{if } x_t - d_t^1 + d_t^2 > \bar{v}_t \text{ and } d_t^2 < s_{Ot}^* \\ x_t - d_t^1 & \text{otherwise.} \end{cases}
$$

The optimal price vectors for the above three cases are $\mathbf{p}_{t+1}^*(x_t - d_t^1 + d_t^2)$, $\mathbf{p}_{t+1}^*(s_{Ot}^*)$ and $\mathbf{p}_{t+1}^*(d_t^2)$;

(III) *If $(x_t, d_t^1, d_t^2) \in \Omega_3$, the net repositioning quantity is given by*

$$
u_t^* = \begin{cases} 0 & \text{if } x_t - d_t^1 + d_t^2 \geq \underline{v}_t \\ x_t - d_t^1 + d_t^2 - s_{It}^* & \text{if } x_t - d_t^1 + d_t^2 < \underline{v}_t \text{ and } d_t^1 < N - s_{It}^* \\ -(N - x_t - d_t^2) & \text{otherwise.} \end{cases}
$$

The optimal price vectors for the above three cases are $\mathbf{p}_{t+1}^*(x_t - d_t^1 + d_t^2)$, $\mathbf{p}_{t+1}^*(s_{It}^*)$ and $\mathbf{p}_{t+1}^*(N - d_t^1)$.

The structural results provide general guidance for the match-back policy in practice (cf. Lam et al., 2007). The idea of match-back policies is intuitive, namely, to maintain the flow conservation at each port, i.e., to equate the container inflow with the outflow. Interestingly, our results suggest that it is not always optimal to maintain this flow conservation. Theorem 3, which is illustrated in Figure 5 for fixed x_t and different combinations of (d_t^1, d_t^2), prescribes when to pursue flow conservation and to what extent it should be maintained.

In particular, it is optimal to not reposition empty containers and thus forgo flow conservation when the state variables fall into region (a) in Figure 5. This happens when the realized demand imbalance is not significant or when both realized demands are so high that no container is idle at either port. As a substitute instrument, the freight rates $\boldsymbol{p}_{t+1}^*$ should be adjusted in accordance with the actual end-of-voyage inventory position $x_t - d_t^1 + d_t^2$. In regions (b) and (c),
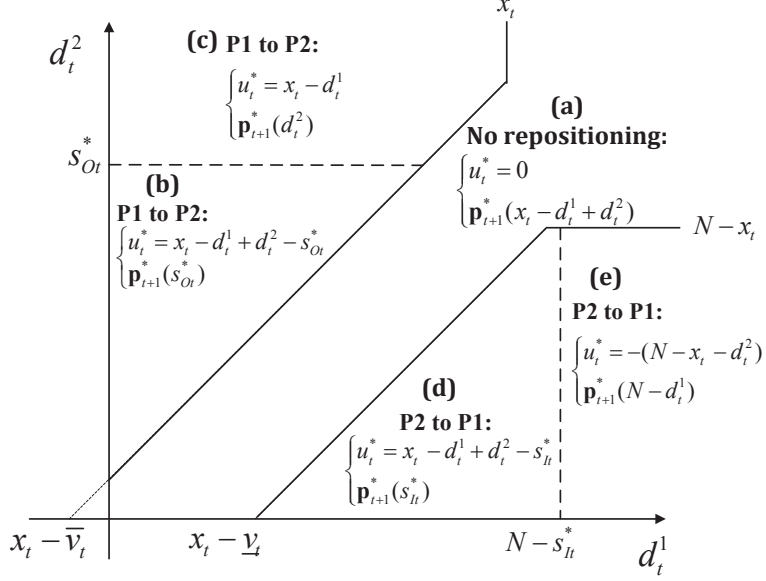
19

Figure 5: The structure of optimal policy (P1: port 1; P2: port 2)

it is optimal to reposition some containers from port 1 to port 2. Our results suggest that there is a target inventory position s_{Ot}^* while repositioning containers, where subscript "O" represents "outbound" from the perspective of port 1. In region (b), the optimal outflow from port 1 equals $d_t^1 + u_t^* = x_t + d_t^2 - s_{Ot}^*$, given an inflow of d_t^2. If the current inventory position is the target level, i.e., $x_t = s_{Ot}^*$, we just need to equate the outflow with the inflow, i.e., $d_t^1 + u_t^* = d_t^2$. Otherwise, we should still maintain flow conservation but leave the end-of-voyage inventory position as s_{Ot}^*. In region (c), demand from port 2 exceeds s_{Ot}^*. It is hence impossible to end up with an inventory position s_{Ot}^*. In this case, we should dispatch all of the empty containers at port 1, and the end-of-voyage inventory position will be d_t^2. Depending on whether or not the target inventory position is achieved, the optimal freight rates should be either $\boldsymbol{p}_{t+1}(s_{Ot}^*)$ or $\boldsymbol{p}_{t+1}(d_t^2)$. Regions (d) and (e) mirror regions (b) and (c) except that port 1 is in deficit and the repositioning direction is reversed. In this case, however, the target inventory position becomes s_{It}^* where the subscript "I" represents "inbound" from the perspective of port 1.

In addition, it is interesting to contrast our results with the well-developed theories in inventory management. As opposed to the celebrated base-stock policy for single-location inventory systems, the allocation of empty containers in our system may oscillate between *two different* inventory positions, depending on the repositioning direction.

Proposition 1. *The target inventory position (at port 1) for outbound repositioning is higher than that for inbound repositioning, i.e., $s_{Ot}^* \geq s_{It}^*$ for all t.*

Proposition 1 highlights another interesting property: The target inventory position is dependent on the repositioning direction. When empty containers are transported from port 1 to port 2,

the target inventory position for port 1 equals s_{Ot}^*; this target level becomes s_{It}^* when empty containers are shipped from port 2 to port 1. Furthermore, $s_{Ot}^* \geq s_{It}^*$. This implies that when trade is unbalanced, it is optimal to aim for a higher inventory position when the port has a capacity surplus than when it has a capacity deficit. Proposition 1 implies that it is *not* economical to maintain the same inventory position for both the deficit and surplus scenarios. This result can be explained by comparing the cost margins for outbound and inbound repositioning. For outbound repositioning, the cost margin for holding one more unit of inventory at port 1 is given by $-c_t^e + h_t^1$, as it increases the inventory holding cost by h_t^1 but reduces the repositioning quantity by one unit. The cost margin is actually negative by Assumption 3, which implies that holding more inventory at port 1 lowers the cost incurred in period t. On the other hand, for inbound repositioning, increasing one more unit of inventory position at port 1 leads to a one-unit increase in the repositioning quantity, and a one-unit decrease in the number of idle containers at port 2. Thus, the cost margin for inbound repositioning is given by $c_t^e - h_t^2$, which is greater than that for outbound repositioning. In addition, no matter whether the repositioning direction is outbound or inbound, the effect of holding one more unit of inventory position at port 1 on future periods (reflected by the function H_t) is the same. Therefore, it is optimal to keep a lower inventory position at port 1 when empty containers are repositioned from port 2 to port 1, as compared to the case when empty containers are repositioned from port 1 to port 2.
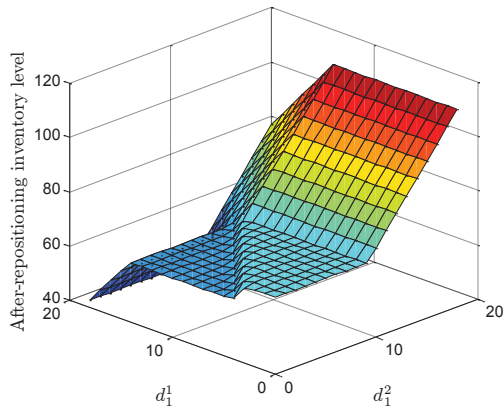

Figure 6: A numerical example of the approximate optimal policy

Figure 6 provides a numerical example[10] to illustrate the end-of-voyage inventory positions in the optimal policy, which validates our analytical result. There are two flat areas in which the inventory positions are constant. Consistent with Proposition 1, the target inventory positions for outbound and inbound repositioning are different from each other.

---

[10]We consider a two-period problem with $N = 150$, $\alpha = 1$, $x_1 = 80$, $p_2^1(\lambda_2^1) = 600 - 4\lambda_2^1$, $p_2^2(\lambda_2^2) = 800 - 4\lambda_2^1$, where $30 \leq \lambda_2^1, \lambda_2^2 \leq 100$. Other parameters are time-invariant and identical for both ports: $c^e = 50$, $c^f = 100$, $h = 30$, $b = 80$, $\epsilon \sim unif[-10, 10]$.

Proposition 2. *For all t, (i) s_{Ot}^* is increasing in c_t^e and decreasing in h_t^1; (ii) s_{It}^* is decreasing in c_t^e and increasing in h_t^2; (iii) consequently, $s_{Ot}^* - s_{It}^*$ is increasing in c_t^e and decreasing in h_t^1 and h_t^2.*

Proposition 2 states how s_{Ot}^* and s_{It}^* change with repositioning and storage costs. When repositioning is more expensive, the gap between s_{Ot}^* and s_{It}^* will be larger as it is less economical to have the same inventory position in both surplus and deficit scenarios. The effect of the holding cost on $s_{Ot}^* - s_{It}^*$ is the opposite, because the gap between the repositioning cost and the holding cost narrows as h_t^i increases.

We close this section by recapping our key analytical findings and their managerial implications. As established in Theorems 1 and 3, in addition to repositioning empty containers, pricing serves as another instrument to cope with demand imbalance. For a port with more excess containers, the price of its outward voyage would be lower to attract more demand. This result is consistent with the empirical finding that for European countries that normally have more imports than exports, inward freight rates are on average 23% higher than for outward ones (cf. De Oliveira, 2014). Additionally, the structure of the optimal repositioning policy characterized in Theorem 3 echoes the match-back policy used in practice (see, e.g., Lam et al., 2007) and we provide conditions under which a simple match back strategy is optimal to our approximate model.

5   Upper Bounds

To evaluate the performance of the control policy generated by the approximate DP, we need to find a relatively tight but computable upper bound of the exact value function $J_1(\mathbf{z}_{0,1}, \mathbf{z}_{1,1}, ..., \mathbf{z}_{L-1,1}, \mathbf{d}_1)$, because solving the exact DP formulation in (6) is extremely time consuming even when L is small due to the high-dimensional state space.

In this section, we adopt the information relaxation-based duality approach developed in Brown et al. (2010) to obtain an upper bound of the exact value function. We consider the perfect information relaxation in which the decision maker in period t is allowed to utilize complete future information, i.e., realizations of random terms $\epsilon_{t+1}, \epsilon_{t+2}, ..., \epsilon_{T+1}$, thus violating the nonanticipativity constraints. Let $\Upsilon = (\hat{\epsilon}_1, \hat{\epsilon}_2, ..., \hat{\epsilon}_T, \hat{\epsilon}_{T+1})$ denote a randomly generated sample path of demand noise. Following Brown et al. (2010), we write the dual of the original DP (6)

as follows:

$$\bar{J}_t(\mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t; \Upsilon) = \max_{\substack{\lambda_{t+1}^i \in [\underline{\lambda}_{t+1}^i, \bar{\lambda}_{t+1}^i], \\ z_{L,t}^j \geq d_t^i \text{ for } i = 1, 2}} \left\{ \alpha R_{t+1}(\boldsymbol{\lambda}_{t+1}) - \sum_{i=1}^2 \left(c_t^e z_{L,t}^i + G_t^i(z_{0,t}^i - z_{L,t}^j) \right) \right.$$

$$-\pi_t(\mathbf{z}_{L,t}, \boldsymbol{\lambda}_{t+1}, \mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t)$$

$$\left. +\alpha \bar{J}_{t+1}(\mathbf{z}_{0,t+1}, ..., \mathbf{z}_{L-1,t+1}, \boldsymbol{\lambda}_{t+1} + \hat{\boldsymbol{\epsilon}}_{t+1,k}; \Upsilon) \right\} + c_t^e(d_t^1 + d_t^2), \tag{17}$$

where

$$\pi_t(\mathbf{z}_{L,t}, \boldsymbol{\lambda}_{t+1}, \mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t) = J_{t+1}^A(\sum_{l=0}^{L-1} z_{l,t+1}^1, \boldsymbol{\lambda}_{t+1} + \hat{\boldsymbol{\epsilon}}_{t+1,k}) - \mathbb{E}_{\boldsymbol{\epsilon}_{t+1}} \left[J_{t+1}^A(\sum_{l=0}^{L-1} z_{l,t+1}^1, \boldsymbol{\lambda}_{t+1} + \boldsymbol{\epsilon}_{t+1}) \right]. \tag{18}$$

In the above DP recursion, the $\mathbf{z}_{l,t+1}$'s are transformed from the $\mathbf{z}_{l,t}$'s according to (3) and (4) as before, but we allow an imaginary decision maker to use the future information on the sample path Υ. The function π_t, defined as the difference between the approximate value function on the sample path and its expected value over $\boldsymbol{\epsilon}_{t+1}$, serves as the penalty function: A strictly positive penalty is imposed whenever the future information brings the imaginary decision maker a higher profit-to-go.[11] We note that π_t is constructed according to Proposition 2.2 in Brown et al. (2010). Therefore, the weak duality holds for problem (17) and the expectation of $\bar{J}_t$ over Υ provides a valid upper bound for the exact value function (6).

Theorem 4. *For all $t = 1, 2, ..., T$ and any state $(\mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t)$,*

$$J_t(\mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t) \leq \mathbb{E}_\Upsilon \left[\bar{J}_t(\mathbf{z}_{0,t}, \mathbf{z}_{1,t}, ..., \mathbf{z}_{L-1,t}, \mathbf{d}_t; \Upsilon) \right]. \tag{19}$$

In the numerical study, we will use simulation to evaluate the expectation in (19), i.e., solve the dual problem (17) on a set of randomly generated sample paths and take the average as the upper bound.

---

[11]If we define $\pi_t = 0$, $\bar{J}_t$ is the value function of our original problem when one can benefit more from future information. Appendix B.1 reports a numerical study that compares the zero-penalty bound with the penalized one according to (18). We note that, strictly speaking, our bound with zero penalty is different from a perfect-information upper bound because the per-period reward is calculated based on the *expected* revenue function $R_t(\boldsymbol{\lambda})$ due to the assumption of additive random noise, rather than the realized revenue on a particular sample path. In some sense, in estimating our upper bounds, we only partially leveraged the future information for minimizing operational costs.

6 Numerical Study

6.1 The Optimality Gap of the Approximation

We quantify the performance of our approximation approach using the upper bound constructed in Section 5. The dynamic programming algorithms for both the approximate model (9) and the dual problem (17) were coded in C++ and compiled with the GNU g++ compiler 6.3.0. The computational experiment in this subsection was conducted on a cluster of Linux workstations where each workstation was equipped with 64GB RAM and a 2.60GHz processor.

First, we test a set of instances with a two-period transit time ($L = 2$). In the numerical experiments with $L = 2$, we discretized the state space into integer values. More precisely, we assume that the inventory level at port 1 within a sufficiently large range is between $-\lceil N/2 \rceil$ and $N + \lceil N/2 \rceil$ and the number of in-transit containers on each vessel is between the lowest possible demand and $\lfloor N/2 \rfloor$.[12] The discretized state space contains all the integer values within these ranges, together with all the possible integer values of demand $\boldsymbol{d}_t$. We consider a stationary setting with $\alpha = 1$, $c_f = 0.5$, $h = 0.01$ and the ϵ_t^i's being independent and identically distributed (iid) according to a uniform distribution $U[-2, 2]$. For notational ease, we suppress the index t whenever appropriate. We vary other parameters as follows: $T \in \{5, 6\}$, $N \in \{25, 30, 35\}$, $b \in \{0.4, 0.6, 0.8\}$, and $c_e \in \{0.2, 0.4, 0.6\}$. The inverse demand function takes the form $A_i(\lambda) = a_i - \lambda$, where we consider two demand settings with $(a_1, a_2) \in \{(14, 12), (16, 10)\}$ to simulate different degrees of demand imbalance. The adjustment range $[\underline{\lambda}_i, \bar{\lambda}_i]$ is set as $[\lambda_i^S - 2, \lambda_i^S + 2]$ where $\lambda_i^S = \frac{a_i - c_f}{2}$, denoting the maximizer of the net revenue $R(\boldsymbol{\lambda}) = \sum_{i=1}^{2}(a_i - \lambda_i - c_f)\lambda_i$. The initial state is set as follows. The number of in-transit containers on each vessel is equal to the optimal container flow in the static problem (8) but the allocation of the idle containers is varied in two different ways: They are (i) equally split between two ports, or (ii) unequally split with one quarter at Port 1 and three quarters at Port 2. We label these two setups as "equal" and "unequal" cases, respectively.

In total, we solved 216 instances with all combinations of the above parameters using our approximation approach. For each instance, we computed the expected profit under the approximate control policy, denoted by J_1^{Approx}, and also evaluated the upper bound, denoted by J_1^{UB}, by solving the dual problem (17) on six randomly generated sample paths. The optimality gap is calculated as $\frac{J_1^{UB} - J_1^{Approx}}{J_1^{UB}}$. Table 1 reports a summary of optimality gaps for different combinations of T, N and initial states, where each combination includes 18 instances. Overall, the average gap is less than 2%. It is worth noting that with the help of monotone properties

---

[12]Recall that the inventory level at port 2 need not be included as a state variable as the total number of containers is fixed. Additionally, we note that $\lceil x \rceil$ ($\lfloor x \rfloor$) denote rounding up (down) x to the nearest integer.

Table 1: Performance of the approximate control policy when $L = 2$

T	N	Initial State	Optimality Gap (%)		
			Average	Median	Max
5	25	Equal	1.88	1.68	3.12
		Unequal	1.96	1.84	3.42
5	30	Equal	0.85	0.91	1.41
		Unequal	1.47	1.47	2.36
5	35	Equal	1.12	0.71	3.27
		Unequal	0.98	0.96	1.56
6	25	Equal	2.24	2.14	3.80
		Unequal	2.59	2.49	5.09
6	30	Equal	1.04	1.03	1.91
		Unequal	2.30	2.28	4.08
6	35	Equal	1.73	1.20	4.79
		Unequal	1.13	0.99	2.28

derived in Section 4.2.2, our approximate DP can be solved within a few seconds. On the other hand, in the case of $T = 6$, depending on the value of N, it can take more than ten hours obtain the upper bound for a single instance. This suggests that even with $L = 2$, it is impossible to solve the exact model within reasonable time, since demand uncertainty would further increase the size of the exact model exponentially over periods.

Table 2: Performance of the approximate control policy when $L = 3$

T	N	Initial State	Conservative Opt. Gap (%)		
			Average	Median	Max
5	30	Equal	4.19	4.08	10.17
		Unequal	4.68	2.89	11.46
5	35	Equal	6.52	5.23	8.34
		Unequal	5.32	5.24	10.78
5	40	Equal	4.18	7.47	12.23
		Unequal	3.60	8.19	14.03
6	30	Equal	3.94	5.04	12.04
		Unequal	4.67	5.12	16.78
6	35	Equal	7.26	3.85	13.24
		Unequal	6.14	3.41	16.29
6	40	Equal	4.12	2.18	14.53
		Unequal	3.15	1.53	15.22

In the second group of experiments, we tested the instances with a three-period transit time ($L = 3$). To emulate a longer transit time, we consider $N \in \{30, 35, 40\}$ and $(a_1, a_2) = \{(10, 8), (12, 6)\}$ to have more containers available relative to the per-period demand. Unfortunately, the upper bound becomes much harder to compute due to the exponentially growing state

space. To estimate the optimality gap within a reasonable time, we further discretized the ranges of the inventory level and the numbers of in-transit containers into 20 and 5 equidistant points, respectively.[13] The demand dimension still contains all possible integer values. Tables 5 and 6 in Appendix B report a numerical study suggesting that such discretization has little impact on the solution quality but would overestimate the optimality gap. As such, our numerical experiments for $L = 3$, conducted under the further discretized state space, provides a rather conservative estimation of the optimality gap. As shown in Table 2, despite being overestimated, the average and the median of gaps are still reasonably small, and the overall average gap is around 4.8%.[14]

Finally, we emphasize that we further discretized the state space merely because we need to compute upper bounds and consistently estimate optimality gaps. It is not necessary to do so when one is interested only in the approximate optimal policy, as the state dimension of our approximate model is independent of L and our approximate model can in fact be applied for even larger L.[15] Unfortunately, we are not able to estimate the optimality gaps for $L \geq 4$ since the upper bound is not computable in that case. To test the performance of our approximate approach with $L \geq 4$, more effective way to evaluate the upper bound may need to be developed and we would leave it for future research.

6.2 The Value of Integrated Decision Making

In this subsection, we investigate the value of coordinating pricing and empty repositioning decisions in two sets of experiments with stationary and time-variant demands, respective. In this subsection, the control policies and their performances are all computed under the state discretization with all integer points.

Stationary Demand. We first focus on a short planning horizon ($T = 4$) with stationary demand, to thoroughly explore how the cost and demand parameters influence the value of coordinating decisions. We assume $L = 1$ such that our approximation is equivalent to the exact model. The other parameters are set as follows. Let $N = 40$ and $\alpha = 1$. Cost parameters are time-invariant and identical for two ports: $c^f = 10$, $h = 4$, $b = 15$, $c^e = 12$. The inverse demand function takes a linear form: $A_i(\lambda) = a_i - \lambda$. We define $\Delta_a = a_1 - a_2$ to capture the degree of potential demand imbalance. In the experiment, we will vary Δ_a while fixing the total market size $a_1 + a_2 = 80$. Random noises ϵ_t^i are iid according to a truncated normal distribution $N(0, \sigma)$

---

[13]In computing the state transition, we chose the nearest point as the next-period state.

[14]Due to the complexity in computing upper bounds, the gap in each instance was estimated on a limited number of sample paths. As such, the worst-case performance was largely influenced by unfavorable sample paths.

[15]Recall that the original problem with L-period transit times has a state space of $2L + 1$ dimensions. By contrast, the dimension of the state space of our approximate model is always three, independent of the specific value of L. We can therefore conclude that the computational time for our approximate model will be generally invariant as L increases.

over $[-5, 5]$, where $\sigma = 10$. The initial state is set as $x_1 = 20$ and $\mathbf{d}_1 = (10, 20)$. As a benchmark, suppose that the liner separates the management of empty containers and pricing decisions, and set λ_i as λ_i^S for all i. Recall that $\boldsymbol{\lambda}^S$ denotes the maximizer of the net revenue $R(\boldsymbol{\lambda})$. In each period, the liner dynamically controls repositioning quantities while fixing $\boldsymbol{\lambda}^S$. We compute the resulting expected profit, denoted by J^S where the superscript S denotes "separate" decision making.
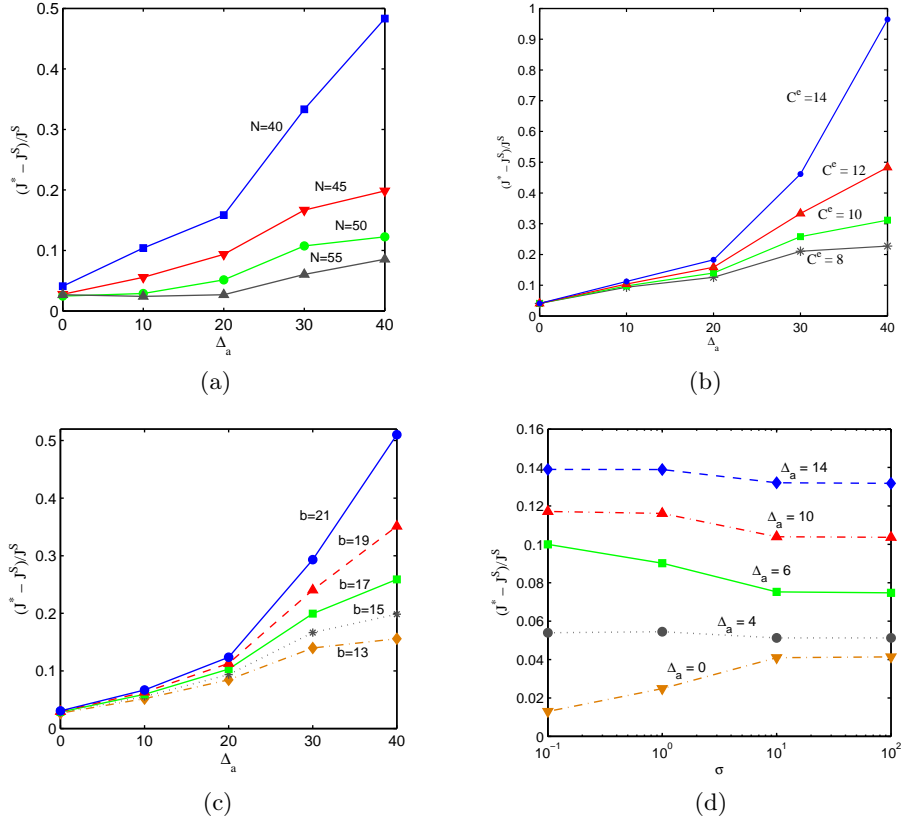

Figure 7: The value of joint decision making

We then allow the decision maker to dynamically adjust λ_i within $[\lambda_i^S - 2, \lambda_i^S + 2]$, coordinating with the empty repositioning. The expected profit is represented by J^*. We quantify the value of integrated decision making by the percentage of improvement, $\frac{J^* - J^S}{J^S}$. The percentage is computed under different Δ_a, N, c^e, b, σ. As shown in Figures 7a, 7b and 7c, the integrated decision making brings greater value as potential demand imbalance Δ_a increases. Moreover, our results indicate that the value in coping with demand imbalance is amplified when the liner owns fewer containers, when handling one additional empty container entails a higher cost c^e, or when the leasing rate for short-term containers b is higher. Nevertheless, the impact of demand uncertainty is ambiguous and depends on the level of Δ_a. It is often observed in single-location inventory control problems that dynamic pricing yields greater improvement over static pricing when the demand is more volatile (e.g., Federgruen and Heching, 1999). In our

two-depot shipping system, however, this is only the case when Δ_a is close to zero (see Figure 7d). Given a moderate demand imbalance ($\Delta_a = 6, 10$), demand volatility in fact offsets the benefit of dynamic pricing for balancing container flows. When demand imbalance is substantial ($\Delta_a \geq 14$), the effect of dynamic pricing on mitigating flow imbalance will become dominant, and demand volatility will then have little impact on profit improvement.

Time-variant Demand. We examine a long planning horizon $T = 52$ with two-period transit times ($L = 2$) to explore the effects of discounting factor α and demand seasonality. One may view it as an annual planning with 52 weeks. We assume that the market alternates between high and low seasons, each season consisting of 13 periods. The inverse demand function in period t is constructed as follows.

$$A_t^i(\lambda) = \begin{cases} (1 + \beta)a_i - \lambda & \text{if period t is in a high season} \\ (1 - \beta)a_i - \lambda & \text{if period t is in a low season} \end{cases}$$

where the a_i's can be interpreted as the baseline market potential and $\beta \in [0, 1]$ is a seasonal factor. Let $\boldsymbol{\lambda}_t^S = (\lambda_t^{1S}, \lambda_t^{2S})$ denote the maximizer of $R_t(\lambda)$, which is time-dependent due to time-variant demand. For integrated decision making, we assume that the adjustment range of $\boldsymbol{\lambda}_t$ is proportional to the market potential. That is, $[\underline{\lambda}_t^i, \bar{\lambda}_t^i]$ is set as $[\lambda_t^{iS} - 2(1 + \beta), \lambda_t^{iS} + 2(1 + \beta)]$ for each high season and as $[\lambda_t^{iS} - 2(1 - \beta), \lambda_t^{iS} + 2(1 - \beta)]$ for each low season. As in the stationary demand setting above, we set $N = 40$ and scale down the demand and cost parameters to accommodate two-period transit times: The baseline market potential is such that $a_1 + a_2 = 40$, $c^f = 5$ and $h = 2$. Random noises ϵ_t^i are iid according to $U[-2, 2]$. We fix $b = 10$, $\Delta_a = 8$ but varies $c^e \in \{7, 6, 5\}$, $\alpha \in \{1, 0.8, 0.6, 0.4\}$, $\beta \in \{0.1, 0.2, 0.3\}$. The initial state is set as in the "equal" case described in Section 6.1. Despite the long planning horizon, our approximate optimal control policy can be found in about 20 minutes.

Table 3: Effects of discounting factor (α) and demand seasonality (β) on the average value of integrated decision making

α	0.4	0.6	0.8	1
$\frac{J^* - J^S}{J^S}$(%)	34.2	37.7	49.7	55.4
β	0.1	0.2	0.3	
$\frac{J^* - J^S}{J^S}$(%)	37.1	47.1	48.5	

Table 3 summarizes the average values of $\frac{J^* - J^S}{J^S}$ in the tested instances with different α and β, where J^* and J^S are the actual profits generated by our approximate optimal policy

and the separate decision making, respectively.[16] On average, our joint optimization model generates greater value (relative to the separate decision making), as the discounting factor α increases. Intuitively, a larger α places greater weights on future rewards and so our model, whereby pricing decisions are forward-looking in consideration of future container allocation, becomes more attractive. In addition, we observed that on average, the value of integrated decision making tends to be higher as demand seasonality increases.

7 Extensions

7.1 Dependent Demand

Our analysis does not rely on the independence of ϵ_t^1 and ϵ_t^2. In each period, the shipping demands for both head-haul and back-haul voyages may have some dependent structure. For instance, during a period of economic prosperity, shipping volumes would be large in both directions. Thus, the results can readily be extended to the case where ϵ_t^1 and ϵ_t^2 are dependent. Furthermore, recall that if a random vector $\boldsymbol{X} = (X_1, X_2)$ is smaller than another random vector $\hat{\boldsymbol{X}} = (\hat{X}_1, \hat{X}_2)$ in the supermodular order, written as $\boldsymbol{X} \leq_{sm} \hat{\boldsymbol{X}}$, then $\hat{X}_1$ and $\hat{X}_2$ are more positively dependent than X_1 and X_2 (Shaked and Shanthikumar, 2007). Proposition 3 proves that when the random noises in opposite directions are more positively dependent, the value of the profit-to-go function (and hence the expected total profit) increases.

Proposition 3. *For all t, if $\boldsymbol{\epsilon}_{t+1} \leq_{sm} \hat{\boldsymbol{\epsilon}}_{t+1}$, i.e., the random vector $\hat{\boldsymbol{\epsilon}}_{t+1} = (\hat{\epsilon}_{t+1}^1, \hat{\epsilon}_{t+1}^2)$ is larger than $\boldsymbol{\epsilon}_{t+1} = (\epsilon_{t+1}^1, \epsilon_{t+1}^2)$ in the supermodular order, then $J_t^A(\mathbf{v}_t; \boldsymbol{\epsilon}_{t+1}) \leq J_t^A(\mathbf{v}_t; \hat{\boldsymbol{\epsilon}}_{t+1})$.*

Intuitively, the positive dependence between the head-haul and back-haul demands reduces the difficulty of balancing container flows to some degree, which enables the liner to obtain a higher expected profit. In addition, demand may also be dependent over time. Our model can be extended to incorporate this dependence by introducing Markov-modulated demand, a widely used modeling technique in the inventory literature (see, e.g., Sethi and Cheng, 1997).

7.2 External Container Flows

In reality, a two-port system can be a sub-system of the entire shipping network operated by an ocean liner. As a result, in addition to the container flows between the two ports, each port may have external inbound or outbound container flows due to other shipping routes. We can model the external flows as a Markov process $\boldsymbol{\xi}_t = (\xi_t^1, \xi_t^2)$, where ξ_t^i denotes the number of

---

[16]We focus on the average value of $\frac{J^* - J^S}{J^S}$ here because $\frac{J^* - J^S}{J^S}$ may not be strictly monotone in α or β when c^e is fixed.

containers arriving and being discharged at port i prior to the t-th voyage. Note that ξ_t^i can also be negative, which indicates that some of empty containers are used (or reserved) for other external service routes. The total number of containers in the two-port system is no longer fixed in the presence of external container flows. Let N_t denote the total number of containers in period t. The approximate DP model can then be adapted to the following:

$$
\begin{aligned}
J_t^A(x_t, \mathbf{d}_t, \boldsymbol{\xi}_t, N_t) = \max_{\lambda_{t+1}^i \in [\underline{\lambda}_{t+1}^i, \bar{\lambda}_{t+1}^i], u_t^i \geq 0} & \{\alpha R_{t+1}(\boldsymbol{\lambda}_{t+1}) - c_t^e \cdot (u_t^1 + u_t^2) \\
& - \hat{G}_t^1(x_t - u_t^1 - d_t^1) - \hat{G}_t^2(N_t - x_t - u_t^2 - d_t^2) + \alpha \mathbb{E}[J_{t+1}^A(x_{t+1}, \mathbf{d}_{t+1}, \boldsymbol{\xi}_{t+1}, N_{t+1}) | \boldsymbol{\xi}_t]\},
\end{aligned}
\tag{20}
$$

where

$$
\begin{aligned}
x_{t+1} &= x_t - u_t^1 - d_t^1 + u_t^2 + d_t^2 + \xi_{t+1}^1, \tag{21} \\
N_{t+1} &= N_t + \xi_{t+1}^1 + \xi_{t+1}^2. \tag{22}
\end{aligned}
$$

For any fixed $(\boldsymbol{\xi}_t, N_t)$, we can apply the same state transformation technique as before. By similar arguments, it can be shown that $J_t^A(\mathbf{v}_t, \boldsymbol{\xi}_t, N_t)$ is $L^\natural$-concave in $\mathbf{v}_t$ for any fixed $(\boldsymbol{\xi}_t, N_t)$. Therefore, the structural properties of the approximate optimal policy still hold when there are external container flows entering or leaving the two-port system. Finally, we note that when the $\boldsymbol{\xi}_t$'s are intertemporally independent, the additional state variables can be reduced from the triplet $(\boldsymbol{\xi}_t, N_t)$ to a single variable N_t.

7.3 Inland Transportation Time

Our model has ignored the distance between the deep-sea terminal and inland customers, and so containers can be immediately used to satisfy the next-period demand upon arrival. In reality, after arriving at a deep-sea terminal, it may take some time for containers to be transferred to inland customers for unloading/loading. To some extent, our framework can be adapted to incorporate the inland transportation time by expanding the transit time. Let us assume that after arrival, it takes one period for each container to be transported to inland depots, unloaded/loaded at nearby customer sites and sent back for another voyage. If the inland transportation time is more than one period, the treatment is analogous.[17]   We can introduce a dummy "in-transit" location before each voyage ends in our model, as illustrated by

---

[17]We note that the inland transportation time can be different for empty and laden containers, as a laden container spends extra time to be discharged at the receiver's place before going for new export cargoes. However, this extra time would be small enough (i.e., less than one period) if the inland transport is well coordinated such that an unloaded container need not travel too far for new cargoes. Hence, we do not distinguish empty and laden containers in accounting the hinterland transportation time.
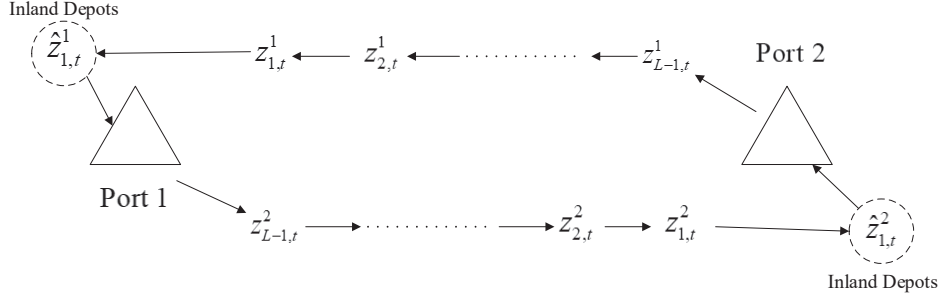
Figure 8: Dynamics of inventory levels

the dotted circles in Figure 8. Each dummy location can represent *multiple* inland container depots in close proximity to customers. Consequently, containers arriving in period t can only be exported (or returned if they are leased) in period $t + 2$ after unloaded and/or loaded at customer sites.This adaptation expands the transit time from L to $L + 1$ periods, and so the proposed approximate model and our analytical results remain valid. However, in doing so, we have assumed that the distance between the deep-sea terminal and each inland depot is similar and that the inland transport is well coordinated. If these assumptions are not satisfied, more sophisticated models and analyses are needed and we leave this for future research.

8 Concluding Remarks

In this paper, we develop a stochastic DP model that integrates pricing decisions with empty container management. The exact DP is intractable due to its high-dimensional state space. Thus, we propose a novel approximation approach whereby the state dimension is reduced to three. Using the properties of $L^{\natural}$-concavity, we characterize the structure of the approximate control policy. An upper bound of the exact DP can be obtained by solving the information relaxation-based dual of the exact DP. Using this upper bound, we numerically show that the control policies generated by our approximate model perform close to optimal when transit times span multiple periods.

Our work has limitations. First, we have restricted to the shipping services with two ports or two regions. In reality, a shipping route may involve more than two ports, which would make the problem more challenging to solve. Since the two-port model studied in this paper has captured the key features of managing empty containers in a cyclic shipping route, we conjecture that our approximation approach can be generalized to more complicated networks with more than two ports. However, characterizing the structural properties would be far more involved. Second, the ocean shipping industry is a competitive market, whereas we only considered a single liner that monopolizes market prices. While Zhou and Lee (2009) and Chen et al. (2016) have developed stylized deterministic models to study the competition between ocean liners, it remains an open

problem to characterize the market equilibrium with uncertain demand.

Acknowledgments

References

Brown, D. B., J. E. Smith, P. Sun. 2010. Information relaxations and duality in stochastic dynamic programs. *Operations research* **58**(4-part-1) 785–801.

Chen, Ming, Zhi-Long Chen. 2015. Recent developments in dynamic pricing research: multiple products, competition, and limited demand information. *Production and Operations Management* **24**(5) 704–731.

Chen, R., J.-X. Dong, C.-Y. Lee. 2016. Pricing and competition in a shipping market with waste shipments and empty container repositioning. *Transportation Research Part B: Methodological* **85** 32–55.

Chen, X., P. Hu, S. He. 2013. Technical note-preservation of supermodularity in parametric optimization problems with nonlattice structures. *Operations Research* **61**(5) 1166–1173.

Chen, X., D. Simchi-Levi. 2004. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research* **52**(6) 887–896.

Chen, X., D. Simchi-Levi. 2012. Pricing and inventory management. *The Oxford handbook of pricing management. Oxford University Press, Oxford* .

Cheung, R. K., C.-Y. Chen. 1998. A two-stage stochastic network model and solution methods for the dynamic empty container allocation problem. *Transportation science* **32**(2) 142–162.

Crainic, T. G., M. Gendreau, P. Dejax. 1993. Dynamic and stochastic models for the allocation of empty containers. *Operations Research* **41**(1) 102–126.

De Oliveira, G. F. 2014. Determinants of european freight rates: The role of market power and trade imbalance. *Transportation Research Part E: Logistics and Transportation Review* **62** 23–33.

Elmaghraby, W., P. Keskinocak. 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science* **49**(10) 1287–1309.

Erera, A.L., J.C. Morales, M. Savelsbergh. 2009. Robust optimization for empty repositioning problems. *Operations Research* **57**(2) 468–483.

Federgruen, A., A. Heching. 1999. Combined pricing and inventory control under uncertainty. *Operations research* **47**(3) 454–475.

Federgruen, A., A. Heching. 2002. Multilocation combined pricing and inventory control. *Manufacturing & Service Operations Management* **4**(4) 275–295.

Fuller, T. 2006. China trade unbalances shipping. *The New York Times* .

Gorman, M. F. 2001. Intermodal pricing model creates a network pricing perspective at bnsf. *Interfaces* **31**(4) 37–49.

Huh, W. T., G. Janakiraman. 2010. On the optimal policy structure in serial inventory systems with lost sales. *Operations Research* **58**(2) 486–491.

King, G. J., H. Topaloglu. 2007. Incorporating the pricing decisions into the dynamic fleet management problem. *Journal of the Operational Research Society* **58**(8) 1065–1074.

Lam, S.-W., L.-H. Lee, L.-C. Tang. 2007. An approximate dynamic programming approach for the empty container allocation problem. *Transportation Research Part C: Emerging Technologies* **15**(4) 265–277.

Lee, C.-Y., M. Yu. 2012. Inbound container storage price competition between the container terminal and a remote container yard. *Flexible Services and Manufacturing Journal* **24**(3) 320–348.

Li, J.-A., S.C.H. Leung, Y. Wu, K. Liu. 2007. Allocation of empty containers between multi-ports. *European Journal of Operational Research* **182**(1) 400–412.

Li, J.-A., K. Liu, S.C.H. Leung, K. K. Lai. 2004. Empty container management in a port with long-run average criterion. *Mathematical and Computer Modelling* **40**(1) 85–100.

Long, Y., L.H. Lee, E.P. Chew. 2012. The sample average approximation method for empty container repositioning with uncertainties. *European Journal of Operational Research* **222**(1) 65–75.

Murota, K. 2009. Recent developments in discrete convex analysis. *Research Trends in Combinatorial Optimization*. Springer, 219–260.

Ng, C. T., D.-P. Song, T.C.E. Cheng. 2012. Optimal policy for inventory transfer between two depots with backlogging. *Automatic Control, IEEE Transactions on* **57**(12) 3247–3252.

Pang, Z., F. Y. Chen, Y. Feng. 2012. Technical note-a note on the structure of joint inventory-pricing control with leadtimes. *Operations Research* **60**(3) 581–587.

Porteus, E. L. 2002. *Foundations of stochastic inventory theory*. Stanford University Press.

Rockafellar, R. T. 1970. *Convex analysis*. Princeton University Press.

Sethi, S. P., F. Cheng. 1997. Optimality of (s, s) policies in inventory models with markovian demand. *Operations Research* **45**(6) 931–939.

Shaked, M., J. G. Shanthikumar. 2007. *Stochastic orders*. Springer.

Song, D.-P., J.-X. Dong. 2015. Empty container repositioning. C.-Y. Lee, Q. Meng, eds., *Handbook of Ocean Container Transport Logistics*. Springer, 163–208.

Song, D.P. 2007. Characterizing optimal empty container reposition policy in periodic-review shuttle service systems. *Journal of the Operational Research Society* **58**(1) 122–133.

Thowsen, G. T. 1975. A dynamic, nonstationary inventory problem for a price/quantity setting firm. *Naval Research Logistics Quarterly* **22**(3) 461–476.

Topaloglu, H., W. Powell. 2007. Incorporating pricing decisions into the stochastic dynamic fleet management problem. *Transportation Science* **41**(3) 281–301.

Topkis, D. M. 1998. *Supermodularity and complementarity*. Princeton University Press.

Zhou, W.-H., C.-Y. Lee. 2009. Pricing and competition in a transportation market with empty equipment repositioning. *Transportation Research Part B: Methodological* **43**(6) 677–691.

Zhu, K., U. W. Thonemann. 2009. Coordination of pricing and inventory control across products. *Naval Research Logistics (NRL)* **56**(2) 175–190.

Zipkin, P. 2008. On the structure of lost-sales inventory models. *Operations Research* **56**(4) 937–944.

Appendices

A $L^{\natural}$-Concavity

Definition 1 ($L^{\natural}$-Concavity)**.** Let $\mathcal{V}$ be a closed convex sublattice of $\mathbb{R}^n$. A function $f : \mathcal{V} \to \mathbb{R}$ is $L^{\natural}$-concave if the function $\psi(\mathbf{v}, \zeta) = f(\mathbf{v} - \zeta \mathbf{e})$ where $\zeta \geq 0$ is supermodular on $\{(\mathbf{v}, \zeta) : \mathbf{v} \in \mathcal{V}, \zeta \in \mathbb{R}_+, \mathbf{v} - \zeta \mathbf{e} \in \mathcal{V}\}$.

Lemma A.1 (Pang et al. (2012))**.** *(i) If $f(\mathbf{v})$ is $L^{\natural}$-concave, so is $\psi(\mathbf{v}, \zeta) = f(\mathbf{v} - \zeta \mathbf{e})$. (ii) If $g(\mathbf{v}, \zeta)$ is $L^{\natural}$-concave, so is $f(\mathbf{v}) = \max_{\zeta \geq 0}\{g(\mathbf{v}, \zeta)\}$. (iii) If $g(\mathbf{v}, \zeta)$ is $L^{\natural}$-concave, and $\zeta(\mathbf{v})$ denotes the smallest value of $\zeta \geq 0$ that maximizes $g(\mathbf{v}, \zeta)$, then $\zeta(\mathbf{v})$ is nondecreasing in $\mathbf{v}$ but $\zeta(\mathbf{v} + \omega \mathbf{e}) \leq \zeta(\mathbf{v}) + \omega$ for $\omega > 0$.*

In addition, like supermodularity, $L^{\natural}$-concavity is preserved under positive linear combination and expectation. Basically, $L^{\natural}$-concavity implies ordinary concavity and supermodularity. As for $L^{\natural}$-convexity, one can find its properties and applications in inventory problems in Zipkin (2008) and Huh and Janakiraman (2010). For more discussions on these concepts, we refer interested readers to Murota (2009) and the literature therein. The following lemma is simple, but will be used frequently in the analysis in this paper.

Lemma A.2. *Let $f(x) : \mathbb{R} \to \mathbb{R}$ be a single variable concave function and (y, z) be a vector from a sublattice of $\mathbb{R}^2$. Then, $f(y - z)$ is $L^{\natural}$-concave in (y, z).*

Proof. By definition, it suffices to show that $\psi(y, z, \zeta) = f[(y - \zeta) - (z - \zeta)] = f(y - z)$ is supermodular in (y, z, ζ). It follows from Lemma 2.6.2(b) in Topkis (1998) that $f(y - z)$ is indeed supermodular in (y, z). $\square$

B Supplementary Numerical Results

B.1 Comparison between Penalized and Non-penalized Upper Bounds

For each instance in the numerical experiments reported in Section 6.1, we computed the upper bound with zero penalty (i.e., $\pi_t = 0$). Table 3 presented a comparison between the bounds with and without the penalty term. On average, imposing a penalty function improves the upper bound. Since the bounds were estimated on only a few simulated sample paths, the zero-penalty bound could be tighter in some instances.

B.2 Impact of the Approximate State Space

We refer to the further discretized state space (i.e., the one containing 20 and 5 discrete values for

Table 4: Penalized Upper Bounds J_1^{UB} versus Non-penalized Upper Bounds J_1^{UB0}

| | T | N | Initial State | $(J_1^{UB} - J_1^{UB0})/J_1^{UB0}$ (%) | | |
				Average	Median	Max
L=2	5	25	Equal	0.00	-0.01	0.25
			Unequal	-0.03	-0.03	0.10
	5	30	Equal	0.20	0.19	0.43
			Unequal	0.24	0.24	0.47
	5	35	Equal	0.33	0.31	0.64
			Unequal	0.28	0.26	0.56
	6	25	Equal	-0.03	-0.01	0.04
			Unequal	0.00	-0.03	0.75
	6	30	Equal	0.06	0.06	0.22
			Unequal	0.17	0.15	0.35
	6	35	Equal	0.17	0.14	0.40
			Unequal	0.25	0.25	0.43
L=3	5	30	Equal	0.12	0.42	1.62
			Unequal	0.33	0.55	1.41
	5	35	Equal	0.39	0.55	1.18
			Unequal	0.46	0.60	1.14
	5	40	Equal	0.37	0.35	0.78
			Unequal	0.54	0.57	1.42
	6	30	Equal	0.25	0.71	1.81
			Unequal	0.33	0.66	1.50
	6	35	Equal	0.27	0.29	1.00
			Unequal	0.32	0.38	1.03
	6	40	Equal	0.25	0.29	0.66
			Unequal	0.34	0.36	1.20

the inventory level and in-transit containers on each vessel, respectively) as the "approximate" state space, as opposed to the all-integer state space (i.e., the one that contains all possible integer points). We focus on the case of $L = 2$ such that the upper bound is computable within a reasonable time. In the experiments, we varied T, N, $a_1 - a_2$ and the initial state while fixing $b = 0.6, c_e = 0.4$ and the other parameters the same as in Section 6.1. For each instance and under each discretization scheme, we computed the approximate optimal policy, and the profit when the policy is implemented in the all-integer state space, and the optimal gap estimated under the corresponding discretization scheme. Table 5 compares the performances and computational times of the control policies computed under different state-space discretization schemes, and Table 6 shows the impact of different state-space discretizations on the estimation of optimality gaps.[18]

---

[18]In Table 5, we use J^{Approx} to denote the profit when the solution obtained in the further discretized state space is implemented in the all-integer state space, and $J^{All-Int}$ to denote the profit when the solution is obtained and also implemented in the all-integer state space as a benchmark. In Table 6, the "Time" columns report the computational time for estimating the upper bounds on five randomly generated sample path.

Table 5: Impact of the approximate state space on the solution quality

N	$a_1 - a_2$	Initial State	Approx.		All-Integer		$\frac{J^{Approx}}{J^{All-Int}}$ (%)
			Profit J^{Approx}	Time (s)	Profit $J^{All-Int}$	Time (s)	
25	2	Equal	457.90	9.42	458.54	18.84	99.86
		Unequal	451.36	9.42	453.70	18.89	99.48
25	6	Equal	468.82	9.37	470.73	18.75	99.59
		Unequal	461.95	9.38	465.93	18.78	99.15
30	2	Equal	465.01	10.25	465.31	24.89	99.94
		Unequal	457.77	10.31	458.13	25.04	99.92
30	6	Equal	478.07	10.23	482.16	24.80	99.15
		Unequal	472.05	10.23	476.14	24.93	99.14
35	2	Equal	468.09	10.99	468.62	30.69	99.89
		Unequal	459.76	10.99	461.08	30.67	99.71
35	6	Equal	486.13	10.98	486.47	30.65	99.93
		Unequal	476.60	11.02	477.86	30.63	99.74

Table 6: Impact of the approximate state space on the estimation of optimality gaps

N	$a_1 - a_2$	Initial State	Approx.		All-Integer	
			Opt. Gap (%)	Time (s)	Opt. Gap (%)	Time (s)
25	2	Equal	0.79	1068.6	0.33	12167.7
		Unequal	1.18	1070.8	0.40	12160.9
25	6	Equal	3.24	903.3	2.27	11890.4
		Unequal	3.40	897.6	2.29	11880.4
30	2	Equal	1.56	1601.0	0.33	34621.5
		Unequal	2.02	1599.0	0.62	34563.8
30	6	Equal	2.25	1408.9	1.26	34819.2
		Unequal	2.60	1409.6	1.42	34815.6
35	2	Equal	0.32	1985.3	0.00	61765.0
		Unequal	1.02	1980.1	0.46	61409.6
35	6	Equal	1.64	1784.0	0.91	62274.9
		Unequal	2.20	1782.3	0.08	62322.6
		Average:	1.85	1457.5	0.86	36224.3

C Technical Lemma

Lemma C.1. *For $t = 1, 2, ..., T$, f_t^A is jointly concave, and has a finite maximizer $(\mathbf{u}_t^*, \boldsymbol{\lambda}_{t+1}^*)$. Moreover, $u_t^{1*} u_t^{2*} = 0$.*

Proof of Lemma C.1. Clearly, J_{T+1}^A is jointly concave as $\hat{G}_{T+1}^i$ is convex for all i by definition. Then, it is straightforward to inductively argue that f_t^A is jointly concave: If J_{t+1}^A is jointly concave, so is J_t^A. This follows since for all t, the feasible region for $(\mathbf{u}_t, \boldsymbol{\lambda}_{t+1})$ is a convex set, and the concavity is preserved under maximization over a convex set by Theorem A.4 in Porteus (2002).

We first argue that at most one of u_t^1 and u_t^2 is nonzero. Suppose in the optimal solution we have $u_t^1 \geq u_t^2 > 0$. Let $\Delta = u_t^1 - u_t^2$. With a fixed $\boldsymbol{\lambda}_{t+1}$, let $\hat{u}_t^1 = \Delta$ and $\hat{u}_t^2 = 0$. In view of (10), x_{t+1} remains the same with the solution $\hat{\mathbf{u}}_t$, and so does $\mathbb{E}J_{t+1}^A(x_{t+1}, \mathbf{d}_{t+1})$. However, by Assumption 3, the cost terms in (9) are reduced at least by $(2c_t^e - h_t^1 - h_t^2)u_t^2 > 0$. This contradicts the optimality of (u_t^1, u_t^2).

To establish the existence of the finite maximizer, we show that given any feasible $\boldsymbol{\lambda}_{t+1}$ and $u_t^j = 0$, $\lim_{u_t^i \to \infty} f_t^A(\mathbf{u}_t, \boldsymbol{\lambda}_{t+1}, x_t, \mathbf{d}_t) = -\infty$. First, note that $J_{T+1}^A(x, \mathbf{d}) \to -\infty$ as $|x| \to \infty$. By induction, if $\lim_{|x| \to \infty} J_{t+1}^A(x, \mathbf{d}) = -\infty$, we have $\lim_{|x| \to \infty} J_t^A(x, \mathbf{d}) = -\infty$. Now, for all t and any given state $(x_t, \mathbf{d})$, as one of u_t^i goes to ∞, we have three terms in $f_t^A(\mathbf{u}_t, \boldsymbol{\lambda}_{t+1}, x_t, \mathbf{d}_t)$ tending to negative infinity while the other two are finite. Hence, for any fixed $\boldsymbol{\lambda}_{t+1}$, $f_t^A \to -\infty$ as either u_t^1 or u_t^2 goes to infinity. Together with the concavity, this implies $\mathbf{u}_t^*$ is finite. $\square$

D Proofs of Main Results

Proof of Lemma 1. In view of Lemma A.2, it is easy to verify the $L^\natural$-concavity of $J_{T+1}^A(\mathbf{v})$. By induction, we need to show that if $J_{t+1}^A(\mathbf{v})$ is $L^\natural$-concave in $\mathbf{v}$, so is $J_t^A(\mathbf{v})$.

First, consider the problem (13). If $J_{t+1}^A(\mathbf{v})$ is $L^\natural$-concave, so is $J_{t+1}^A[(0, y_t^1, y_t^3 + \epsilon_{t+1}^2)^T - (\lambda_{t+1}^1 + \epsilon_{t+1}^1)\mathbf{e}]\}$ in $(y_t^1, y_t^3, \lambda_{t+1}^1)$ for any realization of ϵ_{t+1}^i, where we have invoked Lemma A.1(i). Hence, $\mathbb{E}J_{t+1}^A[(0, y_t^1, y_t^3 + \epsilon_{t+1}^2)^T - (\lambda_{t+1}^1 + \epsilon_{t+1}^1)\mathbf{e}]\}$ is also $L^\natural$-concave in $(y_t^1, y_t^3, \lambda_{t+1}^1)$. This implies the joint concavity in $(y_t^1, y_t^3, \lambda_{t+1}^1)$. By Assumption 2, $R_{t+1}(\lambda_{t+1}^1, y_t^3 - y_t^1)$ is also jointly concave in $(y_t^1, y_t^3, \lambda_{t+1}^1)$, as the composition of affine functions preserves the concavity. The set $\{(y_t^1, y_t^3, \lambda_{t+1}^1) : (y_t^3, \lambda_{t+1}^1) \in A(y_t^1)\}$ is convex. By Theorem A.4 in Porteus (2002), $H_t(y_t^1)$ is simply a single-variable concave function.

Next, we consider the problem (15). By definition, it then suffices to show that $J_t^A(\mathbf{v}_t - \zeta \mathbf{e})$ is supermodular in $(\mathbf{v}_t, \zeta)$ for $\zeta \geq 0$. To this end, we first perform the maximization over u_t:
$$J_t^A(\mathbf{v}_t) = \max_{y_t^2 \geq v_t^3}\{W_t(y_t^2, v_t^2) - c_t^e y_t^2 - \hat{G}_t^2(N - y_t^2 + v_t^1)\} + c_t^e v_t^3, \text{ where } W_t(y_t^2, v_t^2) = \max_{u_t^1 \geq 0}\{H_t(y_t^2 - $$

38

$u_t^1) - c_t^e u_t^1 - \hat{G}_t^1(v_t^2 - u_t^1)\} - c_t^e y_t^2 - \hat{G}_t^2(N - y_t^2 + v_t^1)$. We claim that $W_t(y_t^2, v_t^2)$ is $L^\natural$-concave. In view of Lemma A.1(ii), it suffices to show that the function inside the maximization is $L^\natural$-concave. By Lemma A.2, $H_t(y_t^2 - u_t^1)$ and $-\hat{G}_t^1(v_t^2 - u_t^1)$ are $L^\natural$-concave in (u_t^1, y_t^2) and (u_t^1, v_t^2), respectively. Therefore, $H_t(y_t^2 - u_t^1) - c_t^e u_t^1 - \hat{G}_t^1(v_t^2 - u_t^1)$ is $L^\natural$-concave in (u_t^1, y_t^2, v_t^2), and so is $W_t(y_t^2, v_t^2)$ in (y_t^2, v_t^2).

Now, we examine the supermodularity of $J_t^A(\mathbf{v}_t - \zeta \mathbf{e})$.

$$J_t^A(\mathbf{v}_t - \zeta \mathbf{e}) = \max_{y_t^2 \geq v_t^3 - \zeta} \{W_t(y_t^2, v_t^2 - \zeta) - c_t^e y_t^2 - \hat{G}_t^2(N - y_t^2 + v_t^1 - \zeta)\} + c_t^e(v_t^3 - \zeta)$$

$$= \max_{\hat{y}_t^2 \geq v_t^3} \{W_t(\hat{y}_t^2 - \zeta, v_t^2 - \zeta) - c_t^e \hat{y}_t^2 - \hat{G}_t^2(N - \hat{y}_t^2 + v_t^1)\} + c_t^e v_t^3,$$

where we set $\hat{y}_t^2 = y_t^2 + \zeta$ in the last equality. Due to the $L^\natural$-concavity of $W_t(y_t^2, v_t^2)$, $W_t(\hat{y}_t^2 - \zeta, v_t^2 - \zeta)$ is supermodular in $(\hat{y}_t^2, v_t^2, \zeta)$. Lemma A.2 implies that $-\hat{G}_t^2(N - \hat{y}_t^2 + v_t^1)$ is supermodular in $(\hat{y}_t^2, v_t^1)$. Moreover, $\{(\hat{y}_t^2, \mathbf{v}, \zeta) : \hat{y}_t^2 \geq v_t^3, \mathbf{v} \in \mathcal{V}, \zeta \geq 0\}$ forms a lattice. By Theorem 2.7.6 in Topkis (1998), $J_t^A(\mathbf{v}_t - \zeta \mathbf{e})$ is supermodular in $(\mathbf{v}_t, \zeta)$. □

Proof of Theorem 1. Consider the problem (13). Recall that R_{t+1} is defined as two separable concave functions. By Lemma A.2, the first term $R_{t+1}(\lambda_{t+1}^1, y_t^3 - y_t^1)$ is hence $L^\natural$-concave in $(y_t^1, y_t^3, \lambda_{t+1}^1)$. By Lemma A.1(i) and Lemma 1, it follows that the second term $\mathbb{E}J_{t+1}^A[(0, y_t^1, y_t^3 + \epsilon_{t+1}^2)^T - (\lambda_{t+1}^1 + \epsilon_{t+1}^1)\mathbf{e}]$ is also $L^\natural$-concave in $(y_t^1, y_t^3, \lambda_{t+1}^1)$. Therefore, the objective function of problem (13) is $L^\natural$-concave.

From Lemma A.1(iii), the optimal choice of λ_{t+1}^{1*} *for any fixed* y_t^3 satisfies that $\lambda_{t+1}^{1*}(y_t^1, y_t^3)$ is increasing in (y_t^1, y_t^3) and

$$0 \leq \lambda_{t+1}^{1*}(y_t^1 + \omega, y_t^3 + \omega) - \lambda_{t+1}^{1*}(y_t^1, y_t^3) \leq \omega. \tag{23}$$

On the other hand, as the constraint set is lattice and $L^\natural$-concavity is preserved under maximization, the objective function after optimizing over λ_{t+1}^1 is $L^\natural$-concave in (y_t^1, y_t^3). For the optimal choice of y_t^{3*}, Lemma A.1(iii) implies

$$0 \leq y_t^{3*}(y_t^1 + \omega) - y_t^{3*}(y_t^1) \leq \omega. \tag{24}$$

By definition, $y_t^{3*}(y_t^1) = y_t^1 + \lambda_{t+1}^{2*}(y_t^1)$. Thus, the inequality translates to

$$-\omega \leq \lambda_{t+1}^{2*}(y_t^1 + \omega) - \lambda_{t+1}^{2*}(y_t^1) \leq 0.$$

Now consider the optimal choice of $\lambda_{t+1}^{1*}(y_t^1) = \lambda_{t+1}^{1*}(y_t^1, y_t^{3*}(y_t^1))$. By the monotonicity of

$\lambda_{t+1}^{1*}(y_t^1, y_t^3)$ and (24), it follows that

$$\lambda_{t+1}^{1*}(y_t^1 + \omega) = \lambda_{t+1}^{1*}(y_t^1 + \omega, y_t^{3*}(y_t^1 + \omega)) \geq \lambda_{t+1}^{1*}(y_t^1, y_t^{3*}(y_t^1)) = \lambda_{t+1}^{1*}(y_t^1).$$

Likewise,

$$\begin{aligned}
\lambda_{t+1}^{1*}(y_t^1 + \omega) - \lambda_{t+1}^{1*}(y_t^1) =& \lambda_{t+1}^{1*}(y_t^1 + \omega, y_t^{3*}(y_t^1 + \omega)) - \lambda_{t+1}^{1*}(y_t^1, y_t^{3*}(y_t^1)) \\
\leq& \lambda_{t+1}^{1*}(y_t^1 + \omega, y_t^{3*}(y_t^1) + \omega) - \lambda_{t+1}^{1*}(y_t^1, y_t^{3*}(y_t^1)) \leq \omega,
\end{aligned}$$

where the first inequality follows by (24) and the monotonicity of $\lambda_{t+1}^{1*}(y_t^1, y_t^3)$, and the second equality follows by (23).

As p_{t+1}^i is one-to-one mapped from λ_{t+1}^i, the sensitivities of p_{t+1}^i immediately follow. $\qquad\square$

Proof of Theorem 2. The problem (15) can be recast as

$$J_t^A(\mathbf{v}_t) = \max_{u_t^1 \geq 0}\{\tilde{H}_t(v_t^1, v_t^3, u_t^1) - c_t^e u_t^1 - \hat{G}_t^1(v_t^2 - u_t^1)\} + c_t^e v_t^3, \tag{25}$$

where $\tilde{H}_t(v_t^1, v_t^3, u_t^1) = \max_{y_t^2 \geq v_t^3}\{H_t(y_t^2 - u_t^1) - c_t^e y_t^2 - \hat{G}_t^2(N - y_t^2 + v_t^1)\}$. Note that $\tilde{H}_t$ is $L^\natural$-concave in (v_t^1, v_t^3, u_t^1). This can be verified as follows:

$$\begin{aligned}
\tilde{H}_t[(v_t^1, v_t^3, u_t^1)^T - \zeta\mathbf{e}] =& \max_{y_t^2 \geq v_t^3 - \zeta}\{H(y_t^2 - u_t^1 + \zeta) - c_t^e y_t^2 - \hat{G}_t^2(N - y_t^2 + v_t^1 - \zeta)\} \\
=& \max_{\hat{y}_t^2 \geq v_t^3}\{H(\hat{y}_t^2 - u_t^1) - c_t^e \hat{y}_t^2 - \hat{G}_t^2(N - \hat{y}_t^2 + v_t^1)\} + c_t^e\zeta,
\end{aligned}$$

where the last inequality follows by setting $\hat{y}_t^2 = y_t^2 + \zeta$. The objective is supermodular in $(\hat{y}_t^2, v_t^1, v_t^3, u_t^1, \zeta)$ and the feasible region is a lattice. Hence, $\tilde{H}_t[(v_t^1, v_t^3, u_t^1)^T - \zeta\mathbf{e}]$ is supermodular in $(v_t^1, v_t^3, u_t^1, \zeta)$.

Therefore, the objective function of problem (25) is $L^\natural$-concave in $(u_t^1, v_t^1, v_t^2, v_t^3)$. From Lemma A.1(iii), $u_t^{1*}(\mathbf{v}_t)$ is nondecreasing in $\mathbf{v}_t$ and $0 \leq u_t^{1*}(\mathbf{v}_t + \omega\mathbf{e}) - u_t^{1*}(\mathbf{v}_t) \leq \omega$. Translating $0 \leq u_t^{1*}(\mathbf{v}_t + \omega\mathbf{e}) - u_t^{1*}(\mathbf{v}_t) \leq \omega$ with the original state vector, we have $u_t^{1*}(x_t, d_t^1 - \omega, d_t^2) - u_t^{1*}(x_t, d_t^1, d_t^2) \leq \omega$. In addition, as $u_t^{1*}(\mathbf{v})$ is nondecreasing, we have $u_t^{1*}(\mathbf{v}_t) \leq u_t^{1*}(\mathbf{v}_t + (0, 0, \omega)^T) \leq u_t^{1*}(\mathbf{v}_t + (0, \omega, \omega)^T) \leq u_t^{1*}(\mathbf{v}_t + \omega\mathbf{e})$. This translates to $u_t^{1*}(x_t, d_t^1, d_t^2) \leq u_t^{1*}(x_t, d_t^1, d_t^2 + \omega) \leq u_t^{1*}(x_t + \omega, d_t^1, d_t^2) \leq u_t^{1*}(x_t, d_t^1 - \omega, d_t^2)$, which leads to the first three inequalities in part (i).

Assuming $u_t^{2*} = 0$, we have $y_t^2 = v_t^3$ in the problem (15). The problem then becomes

$$J_t^A(\mathbf{v}_t) = \max_{u_t^1 \geq 0}\{H_t(v_t^3 - u_t^1) - c_t^e u_t^1 - \hat{G}_t^1(v_t^2 - u_t^1)\} - \hat{G}_t^2(N - v_t^3 + v_t^1). \tag{26}$$

The optimal choice of u_t^1 only depends on v_t^2 and v_t^3, and the objective function is $L^\natural$-concave in (v_t^2, v_t^3). From Lemma A.1(iii), it immediately follows that for $\omega > 0$,

$$0 \leq u_t^{1*}(v_t^2 + \omega, v_t^3 + \omega) - u_t^{1*}(v_t^2, v_t^3) \leq \omega.$$

The inequalities translate to $0 \leq u_t^{1*}(x_t + \omega, d_t^1, d_t^2) - u_t^{1*}(x_t, d_t^1, d_t^2) = u_t^{1*}(x_t, d_t^1 - \omega, d_t^2) - u_t^{1*}(x_t, d_t^1, d_t^2) \leq \omega$. It can be seen that a one-unit increase in x_t has the same effect u_t^{1*} as a one-unit decrease in d_t^1, and u_t^{1*} will remain unchanged if x_t and d_t^1 vary by the same amount. $\square$

Proof of Theorem 3. *Step 1.* We first show the following lemma which establishes the differentiability of $H_t(x)$.

Lemma D.1. $H_t(x)$ *is continuously differentiable for all* t.

Proof of Lemma D.1. First, we have established that $H_t(x)$ is a concave function. It is known that if a finite function is convex and differentiable, it is also continuously differentiable (see, for example, Section 25 in Rockafellar, 1970). Hence, we only need to show that $H_t(x)$ is differentiable.

Now consider problem (13). By the envelop theorem, the derivative of $H_t(y_t^1)$ exists as long as $\frac{\partial \mathbb{E} J_{t+1}^A[(0,y_t^1,y_t^3 + \epsilon_{t+1}^2)^T - (\lambda_{t+1} + \epsilon_{t+1}^1)\mathbf{e}]}{\partial y_t^1}$ is well defined (because $\frac{\partial R_{t+1}(\lambda_{t+1}^1, y_t^3 - y_t^1)}{\partial y_t^1}$ exists by Assumption 2 and the relevant constraint is linear and hence differentiable in y_t^1).

To show the differentiability of $\mathbb{E} J_{t+1}^A$ in v_{t+1}^2, we will use the following result from Zhu and Thonemann (2009) (see Claim 4 in their proof of Lemma 1): For continuous functions $f(x)$ and $\phi(x)$, if $f(x)$ is continuously differentiable except on a finite set of points, and the derivative, whenever it exists, is bounded, then $\int f(x - \epsilon)\phi(\epsilon)d\epsilon$ is continuously differentiable. This result basically implies that even when $f(x)$ is not differentiable at a finite number of points, the expectation of $f(x - \epsilon)$ over ϵ is differentiable as long as $f(x)$ has a bounded derivative at differentiable points.

Applying this result, it then remains to show that $J_{t+1}^A(\mathbf{v}_{t+1})$ is differentiable in v_{t+1}^2 except on a finite set of points and its derivative, if it exists, is bounded. This holds for J_{T+1}^A, as J_{T+1}^A is differentiable in v_{T+1}^2 except at $v_{T+1}^2 = 0$ and $|\frac{\partial J_{T+1}^A(\mathbf{v}_{T+1})}{\partial v_{T+1}^2}| \leq \max\{h_{T+1}^1, b_{T+1}^1\}$ for $v_{T+1}^2 \neq 0$. For $t = 1, 2..., T$, applying the envelop theorem in the DP recursion (11), we find that the derivative of $J_t^A(\mathbf{v}_t)$ with respective to v_t^2 only depends on the term $-\hat{G}_t^1(v_t^2 - u_t^{1*})$. Thus, $J_t^A(\mathbf{v}_t)$ is differentiable in v_t^2 except at $v_t^2 = u_t^{1*}$, and $|\frac{\partial J_t^A(\mathbf{v}_t)}{\partial v_t^2}| \leq \max\{h_t^1, b_t^1\}$ for $v_t^2 \neq u_t^{1*}$. That is, for all t, $J_t^A(\mathbf{v})$ is differentiable in v_t^2 except on a finite set of points. $\square$

Step 2. We show that according to the repositioning direction, the state space can be divided

into four segments Ω_i, as defined in the main text and illustrated in Figure 4.

By Assumption 4, it clearly follows that $u_t^* = 0$ if $d_t^1 \geq x_t$ and $d_t^2 \geq N - x_t$; $u_t^* \geq 0$ if $d_t^1 < x_t$ and $d_t^2 \geq N - x_t$; $u_t^* \leq 0$ if $d_t^1 \geq x_t$ and $d_t^2 < N - x_t$. In the above three cases, the repositioning direction is simply due to the capacity constraint.

We proceed to scrutinize the remaining case where $d_t^1 < x_t$ and $d_t^2 < N - x_t$. By the concavity of problem (16) and Lemma D.1, it is optimal to reposition nothing if and only if

$$-H_t'(v_t^3) \in \partial_u W_t(0, \mathbf{v}_t), \tag{27}$$

where $-H_t'(v_t^3)$ is the derivative of $H_t(v_t^3 - u_t)$ with respect to u_t at $u_t = 0$ and $\partial_u W_t(0, \mathbf{v}_t)$ represents the subgradient of W_t with respect to u_t at $u_t = 0$. In fact, equation (27) is just the first-order optimality condition for $u_t^* = 0$. When $d_t^1 < x_t$ and $d_t^2 < N - x_t$, the slope of each piece of $W_t(u_t, \mathbf{v}_t)$ is illustrated in Figure 9. Hence, $\partial_u W_t(0, \mathbf{v}_t)$ is given by the interval $[-c_t^e + h_t^2, c_t^e - h_t^1]$, and the optimality condition translates to $u_t^* = 0$ if and only if $-c_t^e + h_t^1 \leq H_t'(v_t^3) \leq c_t^e - h_t^2$. Recall that $H_t'(v)$ is nonincreasing by concavity. We can define $\underline{v}_t$ as the smallest v such that $H_t'(v) \leq c_t^e - h_t^2$, and $\bar{v}_t$ as the largest v such that $H_t'(v) \geq -c_t^e + h_t^1$, where $\underline{v}_t \leq \bar{v}_t$.[19] By the concavity and optimality condition (27), it follows that $u_t^* = 0$ when $\underline{v}_t \leq v_t^3 \leq \bar{v}_t$; $u_t^* \geq 0$ when $v_t^3 > \bar{v}_t$; and $u_t^* \leq 0$ when $v_t^3 < \underline{v}_t$. That is, when $d_t^1 < x_t$ and $d_t^2 < N - x_t$, the repositioning direction depends on the value of $v_t^3 = x_t - d_t^1 + d_t^2$, and two thresholds $\bar{v}_t$ and $\underline{v}_t$.
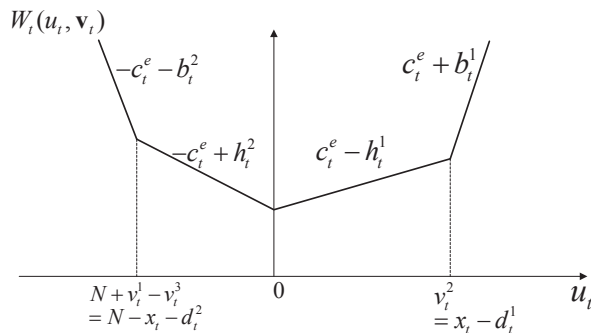

Figure 9: The slopes of $W_t(u_t, \mathbf{v}_t)$ with respect to u_t

Step 3. Having determined the sign of u_t^* in each state segment, we are ready to pin down the specific policies in Ω_i.

For segments Ω_1 and Ω_4, we have shown that the net repositioning quantity $u_t^* = 0$. Therefore, the end-of-voyage inventory position is simply equal to $x_t - d_t^1 + d_t^2$. From Theorem 1, the optimal prices depends only on $x_t - d_t^1 + d_t^2$, written as $\mathbf{p}_{t+1}(x_t - d_t^1 + d_t^2)$.

---

[19]For ease of exposition, in this definition we allow $\bar{v}_t$ and $\underline{v}_t$ to take infinite values. For instance, $\underline{v}_t = -\infty$ (resp. $+\infty$) when $H_t'(v) \leq c_t^e - h_t^2$ (resp. $H_t'(v) > c_t^e - h_t^2$) for all v.

Now we consider Ω_2, where the net repositioning quantity $u_t^* \geq 0$. In view of problem (16) and Assumption 4, the optimal choice of u_t is in effect determined by $\max 0 \leq u_t \leq x_t - d_t^1 \{H_t(v_t^3 - u_t) - (c_t^e - h_t^1)u_t\}$, where we have omitted constant terms in the objective function of (16). Let $y = v_t^3 - u_t$, denoting the target inventory position. The problem is further reduced to

$$\max_{d_t^2 \leq y \leq v_t^3} \{H_t(y) + (c_t^e - h_t^1)y\} - (c_t^e - h_t^1)v_t^3.$$

Let s_{Ot}^* be the (smallest) solution such that $H_t'(s_{Ot}^*) = -c_t^e + h_t^1$, i.e., s_{Ot}^* is the maximizer to the unconstrained problem $\max_y \{H_t(y) + (c_t^e - h_t^1)y\}$. Therefore, we have three subcases in Ω_2:

Subcase 1: $v_t^3 \leq \bar{v}_t$. As $\bar{v}_t$ is the largest value of y such that $H_t'(y) \geq -c_t^e + h_t^1$, by the concavity of H_t, we have $H_t'(v_t^3) \geq -c_t^e + h_t^1$, which implies $s_{Ot}^* \geq v_t^3$. Therefore, the optimal solution is attained at the upper bound of y, i.e, $y^* = v_t^3$.

Subcase 2: $v_t^3 > \bar{v}_t$ and $d_t^2 < s_{Ot}^*$. Similar to Subcase 1, for $v_t^3 > \bar{v}_t$, we have $s_{Ot}^* < v_t^3$. Together with $d_t^2 < s_{Ot}^*$, the interior optimum is attained: $y^* = s_{Ot}^*$. That is, $u_t^* = v_t^3 - s_{Ot}^*$.

Subcase 3: $v_t^3 > \bar{v}_t$ and $d_t^2 \geq s_{Ot}^*$. The optimal solution is attained at the lower bound of y, i.e, $y^* = d_t^2$. That is, $u_t^* = v_t^3 - d_t^2 = x_t - d_t^1$.

The analysis in Ω_3 is analogous. As $u_t \leq 0$, the optimal choice of u_t is essentially determined by $\max_{-(N-x_t-d_t^2) \leq u_t \leq 0} \{H_t(v_t^3 - u_t) + (c_t^e - h_t^2)u_t\}$. Replacing u_t with $y = v_t^3 - u_t$ leads to

$$\max_{v_t^3 \leq y \leq N - d_t^1} \{H_t(y) - (c_t^e - h_t^2)y\} + (c_t^e - h_t^2)v_t^3.$$

Let s_{It}^* denote the (smallest) maximizer to the unconstrained problem $\max_y \{H_t(y) - (c_t^e - h_t^2)y\}$. We have $H_t'(s_{It}^*) = (c_t^e - h_t^2)$. Three subcases need to be considered.

Subcase 1: $v_t^3 \geq \underline{v}_t$. As $\underline{v}_t$ is the smallest value of y such that $H_t'(y) \leq c_t^e - h_t^2$, by concavity of H_t, we have $H_t'(v_t^3) \leq c_t^e - h_t^2$, which implies $s_{Ot}^* \leq v_t^3$. Therefore, the optimal solution is attained at $y^* = v_t^3$, namely, $u_t^* = 0$.

Subcase 2: $v_t^3 < \underline{v}_t$ and $N - d_t^1 > s_{It}^*$. Similar to the discussion in Subcase 1, for $v_t^3 < \underline{v}_t$, we have $s_{Ot}^* > v_t^3$. Together with $N - d_t^1 \geq s_{Ot}^*$, the interior optimum is attained: $y^* = s_{It}^*$. That is, $u_t^* = v_t^3 - s_{It}^*$.

Subcase 3: $v_t^3 > \underline{v}_t$ and $N - d_t^1 \leq s_{It}^*$. The optimal solution is attained at $y^* = N - d_t^1$. That is, $u_t^* = v_t^3 - (N - d_t^1) = -(N - x_t - d_t^2)$.

The corresponding pricing decisions are determined by the values of y^*. □

Proof of Proposition 1. As defined in the proof of Theorem 3, s_{Ot}^* and s_{It}^* are the solutions to unconstrained problems $\max_y \{H_t(y) + (c_t^e - h_t^1)y\}$ and $\max_y \{H_t(y) - (c_t^e - h_t^2)y\}$. This implies that $H_t'(s_{Ot}^*) \leq H_t'(s_{It}^*)$. Hence, $s_{Ot}^* \geq s_{It}^*$ by the concavity of H_t. □

Proof of Proposition 2. The proposition follows immediately from the concavity of H_t and the first-order optimality conditions: $H_t(s_{Ot}^*) + (c_t^e - h_t^1) = 0$ and $H_t(s_{It}^*) - (c_t^e - h_t^2) = 0$. $\square$

Proof of Proposition 3. As we know, if two n-dimensional random vectors $\mathbf{X}$ and $\mathbf{Y}$ satisfy $\mathbf{X} \leq_{sm} \mathbf{Y}$, then $\mathbb{E}[\psi(\mathbf{X})] \leq \mathbb{E}[\psi(\mathbf{Y})]$ for any supermodular function $\psi : \mathbb{R}^n \to \mathbb{R}$. Recall that the transformed DP recursion is given by

$$J_t^A(\mathbf{v}_t; \boldsymbol{\epsilon}_{t+1}) = \max_{(u_t^1, y_t^1, y_t^2, y_t^3, \lambda_{t+1}^1) \in \mathcal{A}} \{\alpha R_{t+1}(\lambda_{t+1}^1, y_t^3 - y_t^1) - c_t^e(u_t^1 + y_t^2) - \hat{G}_t^1(v_t^2 - u_t^1)$$
$$- \hat{G}_t^2(N - y_t^2 + v_t^1) + \alpha \mathbb{E} J_{t+1}^A((0, y_t^1, y_t^3 + \epsilon_{t+1}^2)^T - (\lambda_{t+1}^1 + \epsilon_{t+1}^1)\mathbf{e}) + c_t^e v_t^3.$$

For any fixed $(u_t^1, y_t^1, y_t^2, y_t^3, \lambda_{t+1}^1)$, J_{t+1}^A is supermodular in any realization of $\boldsymbol{\epsilon}_{t+1}$. This follows by the $L^\natural$-concavity of J_{t+1}^A (from Lemma 1). Then we can write $J_t^A(\mathbf{v}_t; \boldsymbol{\epsilon}_{t+1})$ as

$$J_t^A(\mathbf{v}_t; \boldsymbol{\epsilon}_{t+1}) = \mathbb{E}[\psi(u_t^{1*}, y_t^{1*}, y_t^{2*}, y_t^{3*}, \lambda_{t+1}^1, \mathbf{v}_t; \boldsymbol{\epsilon}_{t+1})],$$

where the superscript $*$ indicates the optimal solution given ϵ_{t+1}. It follows that

$$J_t^A(\mathbf{v}_t; \boldsymbol{\epsilon}_{t+1}) = \mathbb{E}[\psi(u_t^{1*}, y_t^{1*}, y_t^{2*}, y_t^{3*}, \lambda_{t+1}^{1*}, \mathbf{v}_t; \boldsymbol{\epsilon}_{t+1})]$$
$$\leq \mathbb{E}[\psi(u_t^{1*}, y_t^{1*}, y_t^{2*}, y_t^{3*}, \lambda_{t+1}^{1*}, \mathbf{v}_t; \hat{\boldsymbol{\epsilon}}_{t+1})]$$
$$\leq \max_{(u_t^1, y_t^1, y_t^2, y_t^3, \lambda_{t+1}^1) \in \mathcal{A}} \mathbb{E}[\psi(u_t^1, y_t^1, y_t^2, y_t^3, \lambda_{t+1}^1, \mathbf{v}_t; \hat{\boldsymbol{\epsilon}}_{t+1})] = J_t^A(\mathbf{v}_t; \hat{\boldsymbol{\epsilon}}_{t+1})$$

$\square$