Approximation Schemes for the Joint Inventory Selection and Online Resource Allocation Problem

Xingxing Chen
Robins School of Business, University of Richmond
Richmond, VA 23173, USA
chen@richmond.edu

Jacob Feldman
Olin Business School, Washington University
St. Louis, MO 63130, USA
jbfeldman@wustl.edu

Seung Hwan Jung
School of Business, Yonsei University
Seoul, 03722 Republic of Korea
seunghwan.jung@yonsei.ac.kr

Panos Kouvelis
Olin Business School, Washington University
St. Louis, MO 63130, USA
kouvelis@wustl.edu

March 29, 2022

Abstract

In this paper, we introduce and study the joint inventory selection and online resource allocation problem, which is characterized by two sequential sets of decisions that are irrevocably linked. First, a decision maker must select starting inventory levels for a set of available resources. Subsequently, the decision maker must match arriving customers to available resources in an online fashion so as to maximize expected reward. We first study the problem in its most general form, before focusing on a specific version that arises at Anheuser Busch InBev (ABI). This particular application of our general setting is referred to as the ABI Trailer Problem, and it considers how ABI ships its beer to vendors via third party delivery trucks. In this problem, ABI must select the weights of preloaded trailers of beer, which are then matched in an online fashion to arriving third party delivery trucks. For the general setting, we develop simple and easy-to-implement approaches that come with robust worst-case performance guarantees. For the ABI setting, we reveal a simplifying structural property related to the optimal matching policy, which gives rise to a natural adaptation of our original approach. We test the efficacy of these policies through extensive numerical experiments, where we find that our approaches are either near-optimal or improve upon state-of-the-art benchmarks. In particular, using a data set from ABI, we are able to generate instances of the ABI Trailer Problem, on which our algorithm has the potential to yield revenue improvements in the range of millions of dollars per year.

1 Introduction

Two problems that have received increased attention in the operations and revenue management literature are those of online resource allocation and inventory selection. The former refers to the problem of allocating/matching customers to a scarce set of resources in an online fashion so as to maximize the reward accrued over a finite selling horizon. There are many widely studied problems that fall within this framework. Examples include the classical network revenue management problem (Talluri and Van Ryzin, 1998; Talluri and van Ryzin, 2004; Zhang and Cooper, 2005), where the goal is to dynamically adjust the set of offered products over a selling horizon to maximize expected revenue when the sale of each product consumes a combination of resources. Other examples include the works of Stein et al. (2020) and Saghafian et al. (2019), who study online matching problems in the context of hospital appointment scheduling. Additionally, the display ads problem (Feldman et al., 2009), which is the edge weighted and capacitated generalization of the online bipartite matching problem, is another widely studied problem in the computer science literature that falls within the framework of online resource allocation.

On the other hand, the inventory selection problem considers how to choose initial inventory levels for a collection of products, which are then subsequently consumed over a finite selling horizon according to some known demand process. In this setting, the decision maker (e.g., retailer) chooses *only* the initial inventory levels and does not make subsequent assortment or matching decisions to adjust for observed demand. Examples of works that have studied this problem include Mahajan and Van Ryzin (2001), Honhon et al. (2010), Goyal et al. (2016), and Aouad et al. (2018), who study different versions of this problem by varying the consumer choice model that governs the demand process.

Interestingly, even though the inventory selection problem is a natural precursor to the online resource allocation problem, there is little work that tackles the two problems simultaneously. For example, Stein et al. (2020) assume that the initial set of available appointment times is exogenously given, and hence only consider the decision of whether to accept or deny incoming requests. Similarly, the many solution approaches developed for the network revenue management problem all assume that the initial inventory level of each resource is fixed and given. In reality, however, inventory levels of resources are often chosen by operational managers. For instance, hospital clinics have control over the set of appointment times and/or rooms to make available for bookings, and airlines control the capacities of each fare class on each flight leg. Moving to the realm of e-commerce, Amazon, for example, decides how to allocate products amongst its many warehouses (the inventory selection problem), and then when orders arrive, they have to decide from which warehouse to satisfy the request (the online matching problem). As such, there are indeed many applications where managers or retailers must first choose initial inventory levels and then couple this decision with a policy used for a subsequent online resource allocation problem.

After initially considering the joint inventory selection and online resource allocation problem in its utmost generality, we introduce and study a lesser-known application, which is a critical piece of the Anheuser Busch Inbev (ABI) supply chain. Specifically, each day, operational managers and brewery warehouse employees at ABI must preload trailers of beer at carefully selected weights, which are then matched to arriving third party delivery trucks, who deliver the beer to wholesalers. In this ABI context, the decision regarding the specific weights of trailers to preload can be viewed as the initial inventory problem, while the subsequent sequence of decisions centered around matching arriving trucks to trailers is an online resource allocation problem. As such, ABI is faced with a problem in which they must develop and link solution approaches to this two-stage problem.

1.1 Problem Formulation

Motivated by the discussion above, we consider a general problem framework in which a decision maker (DM) must select initial inventory levels for a collection of resources, and then decide how to allocate these resources in an online fashion to customers who arrive over a finite time horizon. With respect to the initial inventory decision, we impose various operational constraints that are widely adopted in the literature, and which are formalized shortly. All-in-all, the goal across both sets of decisions is to maximize the cumulative expected reward accrued over the finite time horizon.

Preliminary notation. A DM has access to L resources indexed by the set $\mathcal{L} = \{1, \ldots, L\}$, which arriving customers consume in an online fashion over a finite time horizon consisting of T time periods. At the beginning of the time horizon, the DM must first select starting inventory levels for each resource, which we capture through the inventory vector $X_1 = (x_{1,1}, \ldots, x_{L,1})$, whose ℓ^{th} element $x_{\ell,1} \in \mathbb{Z}_+$ gives the number of units of resource $\ell \in \mathcal{L}$ initially stocked [1]. More generally, we use $X_t \in \mathbb{Z}_+^L$ to denote the remaining inventories of each resource at the beginning of period t, and $\mathcal{L}(X_t) = \{\ell \in \mathcal{L} : x_{\ell,t} > 0\}$ to denote the resources that have yet to stock-out as of period t, and hence can be feasibly matched to arriving customers.

The online matching problem. For a fixed initial inventory vector X_1, we consider the following online matching problem. We assume that each arriving customer belongs to one of B customer types, which we index by the set $\mathcal{B} = \{1, \ldots, B\}$. In each period $t \in [T]$ [2], at most a single customer arrives, and with probability $p_{b,t}$, this customer is of type b. We use $\mathcal{P}_{\text{total}} = \sum_{t \in [T]} \sum_{b \in \mathcal{B}} p_{b,t}$ to denote the expected number of customer arrivals over the time horizon. Each arriving customer must be either deterministically matched to a resource with available inventory, or rejected. If resource l is matched to a customer of type b, then a single unit of resource l is consumed and a reward of $r_{l,b}$ is collected. The structure of the optimal policy for the online matching problem can be derived via the following simple dynamic program, whose value functions $V_t(X_t)$ represent the maximum expected reward that can be garnered from periods $t, \ldots, T$, given that the available inventory at the beginning of period t is X_t. As such, the value function $V_1(X_1)$ gives the maximum

[1] Throughout the paper, we use $\mathbb{Z}_+$ to indicate the set of non-negative integers

[2] For integer $x \in \mathbb{Z}_+$, we let $[x] = \{1, 2, \ldots, x\}$

expected reward that can be derived from a starting inventory of X_1. The Bellman equations for this dynamic program are stated below

$$\begin{aligned} V_t(X_t) &= \sum_{b\in\mathcal{B}} p_{b,t} \cdot \max\left\{ \max_{\ell\in\mathcal{L}(X_t)} \{r_{\ell,b} + V_{t+1}(X_t - e_\ell)\}, V_{t+1}(X_t)\right\} + \left(1 - \sum_{b\in\mathcal{B}} p_{b,t}\right) \cdot V_{t+1}(X_t) \\ &= \sum_{b\in\mathcal{B}} p_{b,t} \cdot \max_{\ell\in\mathcal{L}(X_t)} \left\{[r_{\ell,b} - (V_{t+1}(X_t) - V_{t+1}(X_t - e_\ell))]^+\right\} + V_{t+1}(X_t) \end{aligned} \tag{1}$$

with base cases $V_{T+1}(\cdot) = 0$. We use e_ℓ to denote the unit vector of all zeros, except for a single one in its ℓ^{th} component and $[x]^+ = \max\{x, 0\}$.

The inventory selection problem. Next, we formulate the inventory selection problem that must be considered as a precursor to the online matching problem. We assume that the DM's choice of the initial inventory vector is subject to two operational constraints, which are commonplace in previous works that consider inventory selection problems (Goyal et al., 2016; Aouad et al., 2018; Aouad and Segev, 2019). First, we assume that at most W total units of inventory can be stocked, and so the initial inventory vector must satisfy $\sum_{\ell\in\mathcal{L}} x_{\ell,1} \le W$. The second constraint limits the number of unique resources that can be stocked to C. This constraint is formally stated as $|\mathcal{L}(X_1)| \le C$. We accommodate both inventory constraints through the set $\mathcal{F}(W, C) = \{X_1 \in \mathbb{Z}_+^L : \sum_{\ell\in\mathcal{L}} x_{\ell,1} \le W, |\mathcal{L}(X_1)| \le C\}$, which denotes the set of all feasible starting inventory levels for fixed upper limits W and C. Without loss of generality, we assume that $W \le LT$ and $C \le W$. The former assumption is valid since, when $W \ge LT$, it is clearly optimal to always stock T units of each of the L resources, in which case a stock-out will never occur, and hence any additional inventory is superfluous. The latter assumption follows from noticing that, when $C > W$, we have $\mathcal{F}(W, C) = \mathcal{F}(W, W)$, meaning that the set of feasible starting inventory vectors is unchanged reducing C to W.

With the above-mentioned notation in-hand, we can formulate the initial inventory selection problem as follows

$$Z^*(W, C) = \max_{X_1\in\mathcal{F}(W,C)} V_1(X_1), \tag{2}$$

where $Z^*(W, C)$ gives the maximum expected reward that can be accrued over the time horizon among all starting inventory vectors $X_1 \in \mathcal{F}(W, C)$.

1.2 Contributions

To the best of our knowledge, we are the first to develop algorithms with provable performance guarantees for any sort of joint inventory selection and online resource allocation problem in a multi-period and multi-product setting. In what follows, we summarize our main theoretical results, and also detail the suite of computational experiments used to benchmark the efficacy of our proposed approaches.

The approximation schemes. First, in Section 2.1, we define a simple and easy-to-implement matching policy, in which arriving customers in each period can only be matched to the available resource with the largest matching reward, or rejected. As such, we aptly refer to this policy as the Best-or-Reject (BoR) policy. Next, in Section 2.2 we focus on the so-called "sufficient supply setting", defined over instances with $W \geq \mathcal{P}_{\text{total}}$. Recall that W is the cap on the allowable number of total inventory units initially stocked, and that $\mathcal{P}_{\text{total}}$ is the expected number of arrivals over the time horizon, and hence the sufficient supply setting captures instances where supply exceeds the expected demand. To tackle such instances, we show how to couple our proposed BoR matching policy with an approach to select an initial inventory so as to garner an expected reward no smaller than $\frac{e-1}{4e} \cdot Z^*(\infty, C)$. It is important to note that our performance guarantee in the sufficient supply setting is stated with respect to the more competitive benchmark of $Z^*(\infty, C)$ (rather than $Z^*(W, C)$), which is the optimal expected reward that can be achieved when the DM can stock infinite units of each of the C selected resources. Next, in Section 2.3, we provide an alternative approach for choosing the initial inventory vector, which when coupled with the BoR matching policy of Section 2.1, earns an expected reward of at least $\frac{1}{4} \cdot Z^*(W, C)$, for any $W \in \mathbb{Z}_+$, irrespective of its relationship to $\mathcal{P}_{\text{total}}$. This updated approach, however, requires solving an integer-programming-based fluid approximation, akin to a version of the deterministic linear program in network revenue management with endogenous starting inventory levels.

Computational experiments. In Section 3, we detail an extensive collection of computational experiments aimed at measuring the efficacy of our approach in relation to two benchmark algorithms. More specifically, we first randomly generate a diverse array of problem instances, some of which have $W = \mathcal{P}_{\text{total}}$, and hence fall within the sufficient supply regime, while others have $W \ll \mathcal{P}_{\text{total}}$, which captures settings where inventory is scarce. For each problem instance, we implement an "inventory-adjusted" version of our BoR policy, which we show leads to a strict improvement over the more vanilla version of the BoR policy presented in Section 2.1. For the instances that fall within the sufficient supply framework, we choose initial inventory levels according to the approach outlined in Section 2.2, and for the other instances, the starting inventories are selected based on the approach detailed in Section 2.3. We benchmark our approach against two sophisticated heuristic policies that draw extensively from the revenue management literature. We find that our approach performs at-worst 2% better than both benchmarks across all parameter combinations tested.

The ABI setting. In Section 4, we introduce and study the joint inventory selection and online matching problem faced by ABI, which is henceforth referred to as the ABI Trailer Problem. The distinguishing features of the ABI Trailer Problem are as follows: (i) $W = T = \mathcal{P}_{\text{total}}$, (ii) each arriving truck (customer) must be assigned an available trailer type (resource), and (iii) the matching rewards follow a specific piecewise linear structure. In response to the second problem feature, we implement an adapted version of the inventory-adjusted BoR matching policy mentioned above, in which the option to reject the arriving truck is removed. With this adjustment, our BoR policy reduces precisely to a true greedy policy, which always assigns each arriving truck to the available trailer type with the largest matching reward. While the performance guarantees established in Section 2 no longer hold when rejections are not permitted, we are nonetheless able to provide strong theoretical support for the use of such a policy. Specifically, due to the piecewise linear structure of the matching rewards, we are able to show that the optimal matching decision in each period, and for each truck, can always be reduced to a choice between two specific trailers. Moreover, one of these two trailers, is guaranteed to be the one selected by our myopic greedy policy.

ABI experiments. We also conduct a series of computational experiments, where the performance of our proposed approach is assessed on instances of the ABI Trailer Problem constructed using real historical truck arrival data and estimates of matching rewards from two North American warehouses. The current practice at ABI is to load all trailers at a single weight ($C = 1$), and so via our proposed approach, we are able to measure the potential benefits of utilizing multiple trailer types. Ultimately, our experiments reveal (i) that our approach produces solutions to the ABI Trailer Problem that are within 1% of optimality on average, and (ii) that stocking up to 5 distinct trailer weights can lead to reward improvements of up to 1.3%, which equates to overall annual revenue improvements of millions of dollars. We conclude this section with a discussion regarding how ABI could effectively select the intial set of trailers to preload when they experience stocking costs related to this initial inventory decision.

1.3 Literature Review

In what follows, we summarize four streams of related literature: online matching, network revenue management, inventory selection, and works from a few distinct areas that closely resemble our problem setting.

Online matching. For a fixed collection of resources and inventory levels, the online matching problem considers how to optimally match resources to arriving demand so as to maximize the expected reward over a given time horizon. Variations of this problem have been studied in the online matching literature, where the efficacy of an algorithm is generally measured by its competitive ratio, which compares the performance of the proposed online approach against an optimal offline algorithm that is given access to the sequence of arrivals. For the Display Ads problem, which is the edge weighted and capacitated generalization of the online bipartite matching problem (a generalization of the online matching problem we consider), it is easy to construct simple instances (see Chapter 7 of Mehta (2013)) for which it is not possible to obtain a non-trivial competitive ratios in the adversarial setting. As such, simplifications such as the free disposal model, which relax the capacity restrictions, have led to algorithms that yield a competitive ratio of $\frac{e-1}{e}$ (Feldman et al., 2009). A simplification of this type is not amenable to our setting, as it would require allowing a resource with zero inventory to be assigned to an arriving customer. In a setting where arriving customers can consume multiple units of a particular resource, and where the goal is simply to maximize the number of units consumed, Stein et al. (2020) develop a 0.321-approximation scheme. Wang et al. (2018) consider an online matching setting that is almost identical to ours, and provide an approach that is $(1 - \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{k}} + O(\frac{1}{k}))$-competitive. Chan and Farias (2009) consider a general framework for online matching that they refer to as stochastic depletion problems. In their setting, the matching rewards are a function solely of the resource, and not of the type of the arriving customer. They show that greedy policies achieve at least half of the optimal expected reward under a very mild set of assumptions on the dynamics of the problem.

Network revenue management. Another stream of literature that closely resembles our work is that of approximate techniques for the network revenue management problem. The seminal approach of Simpson (1989) proposes a linear-programming-based approximation of the problem, known as the deterministic linear program (DLP), where the demand for each product is assumed to take on its expected value. Later on, Talluri and Van Ryzin (1998) and Talluri and Ryzin

(1999) study the performance of bid price policies that can be derived from an optimal solution to the DLP. Topaloglu (2009) proposes an alternative way to derive bid prices, which employs a Lagrangian relaxation to decouple decisions across resources. To the best of our knowledge, Ball and Queyranne (2009) are the first to provide performance guarantees for online revenue management problems. They show that no online algorithm can achieve a competitive ratio larger than $\frac{1}{2}$ even for the two fare class single leg revenue management problem. All of the above-mentioned works assume that customer choice is governed by a so-called independent demand model, which fits the profile of the online matching setting that we consider.

Moving to works where demand is governed by some underlying customer choice model, Gallego et al. (2004) propose the choice-based deterministic linear program (CBDLP), where there is a decision variable for the fraction of time to offer each assortment of products over the selling horizon. They show that the optimal objective value of the CBDLP is an upper bound on the optimal expected reward and hence its value can be used to benchmark heuristics. Liu and van Ryzin (2009) extend this work by showing that the optimal objective of the CBDLP converges to the expected reward of an optimal policy as the capacities and length of the time horizon are scaled to infinity. Méndez-Díaz et al. (2010), Gallego et al. (2015) and Feldman and Topaloglu (2017) present approaches for solving the Choice-Based DLP under various popular choice models. Jasin and Kumar (2012) propose an approach that continuously re-solves the CBDLP to reflect changing inventory levels, and they show that such a policy achieves bounded regret.

More recently, the work of Ma et al. (2020) builds on the above-mentioned results to achieve a constant factor performance guarantee for the network revenue management problem in its utmost generality, which can easily be adapted to our general online matching setting. Finally, the works of Gong et al. (2021) and **?** provide constant factor guarantees for online resource allocation problems with reusable resources, i.e. each resource that is consumed becomes available again at a later (random) time period.

Inventory selection. Next, we review works that consider the problem of optimally choosing initial inventory levels for a collection of products, which are subsequently consumed over a finite selling horizon. As noted above, Mahajan and Van Ryzin (2001), Honhon et al. (2010), Goyal et al. (2016), and Aouad et al. (2018) consider a version of this problem in which the retailer cannot vary her assortment for each arriving customer, and hence her only decision consists of choosing an initial inventory level for each product subject to a constraint on the total number of units stocked. The first three works assume that customers make purchasing decisions according to a non-parametric ranking-based choice model, in which each customer is distinguished by a ranking on a subset of products, and will ultimately purchase her highest ranking product that is available. Mahajan and Van Ryzin (2001) consider a single period version of the problem, and show various structural properties related to the expected profit function. To the best of our knowledge, Honhon et al. (2010) is the first to consider a multi-period version of this problem, and they provide an optimal algorithm whose running time scales exponentially in the number of products. Goyal et al. (2016) considers a version of the problem with restricted sets of customer preference rankings, and provide a constant factor guarantee whose running time is polynomial in all input parameters. Aouad et al. (2018) and Aouad and Segev (2019) assume that customer choice is governed by a Multinomial Logit (MNL) choice model; the former work provides a polynomial time constant-factor guarantee, while the more recent work yields a nuanced polynomial time approximation scheme (PTAS). Finally, Martínez-de Albéniz and Kunnumkal (2021) consider a variant of the inventory selection problem under MNL preferences with replenishment. They propose an integer-programming-based approach, which exploits a closed-form solution for the single-product setting.

The most closely related work. To the best of our knowledge, Dong et al. (2009), Gallego and Kim (2020) and El Housni et al. (2021) are the only other works that focus on a problem setting that is close to ours, albeit where demand is modeled via a particular choice model. Furthermore, the problem settings considered in the three aforementioned papers differ from ours in that the initial inventory decision is unconstrained, but the DM incurs a cost for each unit stocked. The first paper considers a pricing setting, and shows that various heuristic approaches are asymptotically optimal as inventories and capacities are scaled to infinity. The latter two provide regret-based analyses of their proposed approaches for both assortment and pricing variants of the problem. Consequently, none of these three works provide approximation algorithms that come with performance guarantees for arbitrary problem instances.

Topaloglu (2013) also considers a closely related problem where a DM must choose initial inventories of products, and then subsequently decide how to vary the assortment of products offered to each arriving customer with the goal of maximizing revenue over a fixed selling horizon. Given that choice is governed by an MNL model, the author shows how to exploit the special structure of the MNL choice probabilities to develop a concise nonlinear programming formulation of the problem, which can be solved efficiently and performs well on practical instances of the problem. However, this approach does not come with any theoretical performance guarantees. Previgliano and Vulcano (2021) consider a related version of this problem in the context of the classic network revenue management problem, where the capacities assigned to each flight leg are selected at some pre-determined time period in the middle of the selling horizon.

Another stream of related work considers Assemble-to-Order (ATO) systems, in which a DM must order parts, and then subsequently use these parts to assemble various products to meet just-revealed demand. Reiman and Wang (2015) provide policies for this problem that are shown to be asymptotically optimal as lead times tend to infinity, whereas DeValve et al. (2020) give a constant factor approximation scheme that uses duality-based arguments.

2 The Approximation Schemes

In this section, we present our approximation schemes for the joint inventory selection and online resource allocation problem. We describe two approaches, each of which makes use of the BoR matching policy presented in Section 2.1, which, for a fixed initial inventory, starts by partitioning the customer types based their highest-reward match, and then utilizes this partition to decompose the problem by resource. We then show how to couple this matching policy with a carefully chosen initial inventory vector for both the sufficient supply (Section 2.2) and general-W (Section 2.3) settings. All proofs for this section can be found in Appendix A.

2.1 The Best-or-Reject Policy for the Online Matching Problem

In order to develop a simple and easy-to-implement policy for the online matching problem, we consider the BoR policy that either rejects the arriving customer, or matches her to her "ideal" resource, i.e. the resource that earns the highest matching reward across all those initially stocked. It is important to note that, in the analysis that follows, we assume that each customer's ideal resource is irrevocably defined with respect to the initial inventory level X_1, and hence does not change as resources are depleted. In Appendix D, we detail a more practically relevant version of this policy, in which the ideal resource of each customer type is updated after each stock-out. We show that this inventory-adjusted BoR matching policy is guaranteed to improve upon the more vanilla version of the BoR policy presented below, although its use does not lead to improved worst-case theoretical performance guarantees.

The BoR value functions. Given the nature of the BoR policy defined above, it will be helpful to first introduce notation to formalize the partitioning of the customer types by their ideal resource. To begin, for a fixed initial inventory vector X_1, let $\ell(b, X_1) = \text{argmax}_{\ell' \in \mathcal{L}(X_1)} r_{\ell',b}$ denote the ideal resource of customer type b. Furthermore, let $\mathcal{B}_\ell(X_1) = \{b \in \mathcal{B} : \ell = \ell(b, X_1)\}$ be the set of customer types whose ideal resource is ℓ, where ties between two resources can be broken arbitrarily. Given that the ideal resources are defined with respect to X_1, let $\hat{V}_t(X_t; X_1)$ represent the maximum expected reward that can be garnered from periods $t, \ldots, T$ under the BoR matching policy described above, when the available inventory at the beginning of period t is X_t. Formally, we have that

$$\hat{V}_t(X_t; X_1) = \sum_{\ell \in \mathcal{L}(X_t)} \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot \left[r_{\ell,b} - (\hat{V}_{t+1}(X_t; X_1) - \hat{V}_{t+1}(X_t - e_\ell; X_1)) \right]^+ + \hat{V}_{t+1}(X_t; X_1), \tag{3}$$

with base cases $V_{T+1}(\cdot; X_1) = 0$. Note that in (3), each arriving customer of type $b \in \mathcal{B}$ must either be assigned to its ideal resource $\ell(b, X_1)$, or rejected. For ease of notation moving forward, we will simply use $\hat{V}_1(X_1)$ to denote $\hat{V}_1(X_1; X_1)$. Next, we show that the value functions defined in (3) admit a simple decoupling by resource, and hence can be computed efficiently.

The decoupling. For each resource $\ell \in \mathcal{L}(X_1)$ and inventory level $x \in [x_{\ell,1}]$, let $\hat{V}_t^\ell(x; X_1)$ represent the maximum expected reward that can be garnered from resource ℓ over periods $t, \ldots, T$, given that its inventory level at the beginning of period t is x, and given that an arriving customer of type b can be matched to resource ℓ if and only if $b \in \mathcal{B}_\ell(X_1)$. We have that

$$\hat{V}_t^\ell(x; X_1) = \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot \left[r_{\ell,b} - (\hat{V}_{t+1}^\ell(x; X_1) - \hat{V}_{t+1}^\ell(x-1; X_1)) \right]^+ + \hat{V}_{t+1}^\ell(x; X_1), \tag{4}$$

with base cases $\hat{V}_{T+1}^\ell(\cdot, \cdot) = 0$ and $\hat{V}_t^\ell(0; X_1) = 0$. Note that this dynamic program has a single-dimensional state space, and hence we can easily compute its value functions through backward recursion. The following claim reveals our decoupling of interest.

Claim 2.1. *For any fixed initial inventory vector X_1, and any period $t \in [T]$, we have*

$$\hat{V}_t(X_t; X_1) = \sum_{\ell \in \mathcal{L}(X_t)} \hat{V}_t^\ell(x_{\ell,t}; X_1)$$

Structural properties. In what follows, we establish two claims related to the optimal value functions given in (4), which will prove critical moving forward. In stating the first of the two claims, we use $\Delta\hat{V}_t^\ell(x;X_1) = \hat{V}_t^\ell(x;X_1) - \hat{V}_t^\ell(x-1;X_1)$ to denote the marginal value of the x^{th} unit of resource ℓ at the beginning of period t.

Claim 2.2. *For an arbitrary starting inventory vector X_1 and an arbitrary resource $\ell \in \mathcal{L}(X_1)$, we have that*

(i) $\hat{V}_t^\ell(x+1;X_1) \geq \hat{V}_t^\ell(x;X_1)$ *for inventory levels* $x \geq 0$,

(ii) $\Delta\hat{V}_t^\ell(x;X_1) \geq \Delta\hat{V}_t^\ell(x+1;X_1)$ *for inventory levels* $x \geq 1$,

(iii) $\Delta\hat{V}_t^\ell(x;X_1) \geq \Delta\hat{V}_{t+1}^\ell(x;X_1)$ *for inventory levels* $x \geq 1$.

Claim 2.3. *For arbitrary starting inventory vectors X_1 and X_1' that satisfy $\mathcal{L}(X_1') \subseteq \mathcal{L}(X_1)$, and any resource $l \in \mathcal{L}(X_1')$ with inventory level x, we have that $\hat{V}_t^\ell(x;X_1') \geq \hat{V}_t^\ell(x;X_1)$.*

The first of the three properties stated in Claim 2.2 implies that more inventory of a particular resource is always more valuable. The second and third properties respectively show that the marginal value of a unit of resource ℓ is decreasing in the inventory level, and increasing in the number of remaining time periods. It is worth noting that these three properties also appear in Talluri and van Ryzin (2004), who consider the choice-based network revenue management problem. We simply adapt their proofs to our matching setting. Claim 2.3 establishes that the value functions associated with each resource increase as its set of greedily assigned customer types grows larger.

2.2 The Sufficient Supply Setting

In this section, we consider instances in which $W \geq \mathcal{P}_{\text{total}}$, and show how to couple the BoR matching policy with a simple approach for selecting the initial inventory vector, so as to garner a constant fraction of $Z^*(\infty, C)$. Our main theorem in this setting is stated below, where we use $\lfloor \cdot \rfloor$ as the standard floor operator that rounds its input down the nearest integer.

Theorem 2.4. *There is a polynomial-time algorithm that computes an initial inventory vector $X_1^{\ll} \in \mathcal{F}(W,C)$ that satisfies $\hat{V}_1(X_1^{\ll}) \geq \frac{e-1}{2e} \cdot \min\{1, \frac{W}{\lfloor \mathcal{P}_{\text{total}} \rfloor + C}\} \cdot Z^*(\infty, C)$.*

It is useful to observe that, since we assume $W \geq \mathcal{P}_{\text{total}}$, we have $\frac{W}{\lfloor \mathcal{P}_{\text{total}} \rfloor + C} \geq \frac{1}{2}$. As a result, $\frac{e-1}{4e} \cdot Z^*(\infty, C)$ is a worst-case performance guarantee of our approach, which is achieved when $W = \mathcal{P}_{\text{total}}$. To give some context to the strength of this result, we note that it is possible to exploit existing inapproximability result to show that, even when $W = \infty$, no approach can garner an expected reward that exceeds $(\frac{e-1}{e}) \cdot Z^*(\infty, C)$, for general $C \in \mathbb{Z}_+$. This result is formalized in Appendix C.1.

Technical overview. We prove Theorem 2.4 over the remainder of this section. To do so, we first establish two lower bounds on the optimal expected reward garnered by the BoR policy, which hold for any choice of starting inventory vector X_1. Then, we exploit the structure of these two lower bounds to show how to choose an initial inventory vector $\bar{X}_1 \in \mathcal{F}(\lfloor \mathcal{P}_{\text{total}} \rfloor + C, C)$ that satisfies $\hat{V}_1(\bar{X}_1) \geq \frac{e-1}{2e} \cdot Z^*(\infty, C)$. If $\sum_{\ell \in \mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} > W$, meaning the starting inventory vector $\bar{X}_1$ is not feasible, we detail how to remove $\sum_{\ell \in \mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} - W$ units of inventory from $\bar{X}_1$ to arrive at $X_1^{\ll} \in \mathcal{F}(W, C)$. Finally, we show that when $X_1^{\ll}$ is coupled with our BoR matching policy, it achieves the performance guarantee stated in Theorem 2.4.

The two lower bounds. For arbitrary starting inventory vector X_1, we use the structural properties laid out in Claim 2.2 to provide two lower bounds on $\hat{V}_1(X_1)$ that are instrumental to proving Theorem 2.4. The first and simpler lower bound is stated in the lemma that follows.

Lemma 2.5 (Lower Bound 1). *For any starting inventory $X_1 \in \mathbb{Z}_+^L$, we have that*

$$\hat{V}_1(X_1) \geq \sum_{\ell \in \mathcal{L}(X_1)} x_{\ell,1} \cdot \Delta \hat{V}_1^{\ell}(x_{\ell,1}; X_1).$$

The second lower bound, which requires a bit more leg-work to establish, is stated below.

Lemma 2.6 (Lower Bound 2). *For any starting inventory $X_1 \in \mathbb{Z}_+^L$, we have that*

$$\hat{V}_1(X_1) \geq \sum_{t \in [T]} \sum_{b \in \mathcal{B}} p_{b,t} \cdot \max_{\ell \in \mathcal{L}(X_1)} r_{\ell,b} - \sum_{\ell \in \mathcal{L}(X_1)} \left(\sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \right) \cdot \Delta \hat{V}_1^{\ell}(x_{\ell,1}; X_1).$$

By combining the two individual lower bounds, we arrive at the following tighter lower bound on the performance of our BoR matching policy for a fixed initial inventory X_1

$$\hat{V}_1(X_1) \geq \max \left\{ \underbrace{\sum_{t \in [T]} \sum_{b \in \mathcal{B}} p_{b,t} \cdot \max_{\ell \in \mathcal{L}(X_1)} r_{\ell,b} - \sum_{\ell \in \mathcal{L}(X_1)} \left(\sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \right) \cdot \Delta \hat{V}_1^{\ell}(x_{\ell,1}; X_1)}_{\text{Lower Bound 2}}, \underbrace{\sum_{\ell \in \mathcal{L}(X_1)} x_{\ell,1} \cdot \Delta \hat{V}_1^{\ell}(x_{\ell,1}; X_1)}_{\text{Lower Bound 1}} \right\}. \tag{5}$$

Since the right-hand-side of (5) is solely a function of the starting inventory decision X_1, it is only natural to focus next on selecting an initial inventory vector to make this right-hand-side as large as possible.

Building $\bar{\boldsymbol{X}}_{\mathbf{1}}$. We first present the following intermediate result, which considers the problem of choosing an initial inventory vector when the DM can stock infinite units of each resource that is initially stocked. We note that this result is a mere consequence of the fact that $\hat{V}_1(X_1)$ is a monotone submodular set function in X_1, when $W = \infty$. As a result, well-known results on submodular maximization can be applied.

Claim 2.7. *There exists a polynomial-time algorithm that constructs an initial inventory vector $X'_C \in \mathcal{F}(\infty, C)$ that satisfies*

$$\hat{V}_1(X'_C) \geq \frac{e-1}{e} \cdot Z^*(\infty, C).$$

Exploiting Claim 2.7, we construct $\bar{X}_1$ from X'_C as follows

$$\bar{x}_{\ell,1} = \begin{cases} \lceil \sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(X'_C)} p_{b,t} \rceil, & \text{if } \ell \in \mathcal{L}(X'_C) \\ 0, & \text{otherwise ,} \end{cases} \tag{6}$$

where $\lceil \cdot \rceil$ is the ceiling operation that rounds its input up to the nearest integer. Intuitively, we build $\bar{X}_1$ to closely resemble X'_C. In other words, we attempt to mirror the inventory vector X'_C by constructing $\bar{X}_1$ using the same resources as X'_C, and then choosing the initial inventory levels of each of these resources so they can always meet their expected demand under our BoR matching policy.

Next, we prove the following lemma, which establishes a lower bound on the expected reward garnered when our BoR matching policy is seeded with $\bar{X}_1$ as the starting inventory.

Lemma 2.8. *With regards to the inventory vector $\bar{X}_1$ as defined in* (6)*, we have that (i) $\bar{X}_1 \in \mathcal{F}(\lfloor \mathcal{P}_{\text{total}} \rfloor + C, C)$ and (ii) $\hat{V}_1(\bar{X}_1) \geq \frac{e-1}{2e} \cdot Z^*(\infty, C)$*

Building $\boldsymbol{X}_{\mathbf{1}}^{\ll}$. The final step is to remove units from $\bar{X}_1$ so as to ensure that at most W total units are stocked. To do so, we first compute $\hat{V}_1^\ell(x; \bar{X}_1)$ for each resource $\ell \in \mathcal{L}(\bar{X}_1)$ and each inventory level $x \in [\bar{x}_{\ell,1}]$. Then, we simply remove the $[\sum_{\ell \in \mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} - W]^+$ units of inventory from $\bar{X}_1$ with the smallest marginal value, as measured by $\Delta \hat{V}_1^\ell(\cdot; \bar{X}_1)$. This idea is formalized in Algorithm 1, which returns $X_1^{\ll}$.

First, we clearly have that $X_1^{\ll} \in \mathcal{F}(W, C)$, since Algorithm 1 continues to execute until $X_1^{\ll}$ stocks W units of inventory. Our final steps towards proving Theorem 2.4 are to establish the desired performance guarantee, and to analyze the running time needed to compute $X_1^{\ll}$, the former of which is accomplished via the following lemma, while the latter is subsequently discussed.

Algorithm 1 Building $X_1^{\ll}$

procedure GreedyRemove($\bar{X}_1$)
    Compute $\Delta\hat{V}_1^{\ell}(x; \bar{X}_1)\ \forall \ell \in \mathcal{L}(\bar{X}_1), x \in [\bar{x}_{\ell,1}]$
    $X_1^{\ll} \leftarrow \bar{X}_1$
    while $\sum_{l \in \mathcal{L}(X_1^{\ll})} x_{l,1}^{\ll} > W$ **do**
        $l^* \leftarrow \operatorname{argmin}_{l \in \mathcal{L}(X_1^{\ll})} \Delta\hat{V}_1^{\ell}(x_{\ell,1}^{\ll}; \bar{X}_1)$
        $X_1^{\ll} \leftarrow X_1^{\ll} - e_{l^*}$
    end while
    return $X_1^{\ll}$
end procedure

Lemma 2.9. *We have*

$$\hat{V}_1(X_1^{\ll}) \geq \frac{e-1}{2e} \cdot \min\{1, \frac{W}{\lfloor \mathcal{P}_{\text{total}} \rfloor + C}\} \cdot Z^*(\infty, C).$$

The final running time. To conclude the proof of Theorem 2.4, we analyze the running time required to compute $X_1^{\ll}$. To do so, we list the three steps required to compute $X_1^{\ll}$, and argue that each step can be executed in a running time that is polynomial in the input.

- Step 1: Compute X'_C as defined in Claim 2.7. As noted in this claim, this starting inventory vector can be derived in polynomial time using existing approaches.

- Step 2: Compute $\bar{X}_1$ as defined in (6). For this step, the bottleneck in terms of computation time is computing $\mathcal{B}_\ell(\bar{X}_1)$ for each $\ell \in \mathcal{L}(\bar{X}_1)$. Recalling that $\mathcal{B}_\ell(X_1)$ is the set of customer types whose ideal resource is ℓ, we note that by simply enumerating over each customer type-resource pair, we can construct $\{\mathcal{B}_\ell(\bar{X}_1)\}_{\ell \in \mathcal{L}(\bar{X}_1)}$ in a running time of $O(LB)$.

- Step 3: Run Algorithm 1, whose running-time-bottleneck is the computation of $\Delta\hat{V}_1^{\ell}(x; \bar{X}_1)$ for each $\ell \in \mathcal{L}(\bar{X}_1)$ and $x \in [\bar{x}_{\ell,1}]$. Noting that $\bar{x}_{\ell,1} \leq W \leq LT$, we see that there are at most $O(L^2T)$ such marginal values, each of which can be computed in polynomial time via the decoupled value function given in (4).

2.3 An Extension for General Inventory Constraints

In this section, we provide an alternative approach for choosing the initial inventory vector, which when coupled with the BoR matching policy, earns an expected reward of $\frac{1}{4} \cdot Z^*(W, C)$, for any $W \in \mathbb{Z}_+$. The caveat, is that this new approach requires solving a simple integer program, and hence its theoretical worst-case running time is not polynomial in the input. That said, in Section 3, we demonstrate that this approach admits an efficient implementation, while also outperforming state-of-art benchmarks in the process.

Technical overview. Our new approach begins by considering a fluid version of our problem, formulated as an integer program (Fluid-IP), where demand takes on its expected value. From the optimal solution to Fluid-IP, we once again partition the customer types based on their ideal resource, among those that are stocked. The total amount of "fluid" demand assigned to each resource is then used to choose the starting inventory levels. From here, we then employ the BoR matching policy, showing in a similar fashion to the analysis of Section 2.2, that this policy achieves the desired performance guarantee.

The fluid integer program. In what follows, we present a standard fluid approximation of our joint inventory and online matching problem akin to ever-popular deterministic linear program in the revenue management literature (Talluri and Van Ryzin, 1998; Liu and van Ryzin, 2009), albeit with endogenous initial inventory levels.

$$
\begin{aligned}
\mathrm{Fluid}(W, C) = \max \; & \sum_{t \in [T]} \sum_{\ell \in \mathcal{L}} \sum_{b \in \mathcal{B}} r_{\ell,b} y^t_{\ell,b} && \text{(Fluid-IP)} \\
\text{s.t.} \; & \sum_{\ell \in \mathcal{L}} y^t_{\ell,b} \leq p_{b,t} && \forall b \in \mathcal{B}, \forall t \in [T] \\
& \sum_{t \in [T]} \sum_{b \in \mathcal{B}} y^t_{\ell,b} \leq x_\ell && \forall \ell \in \mathcal{L} \\
& \sum_{\ell \in \mathcal{L}} x_\ell \leq W && \forall \ell \in \mathcal{L} \\
& x_\ell \leq W \cdot z_\ell && \forall \ell \in \mathcal{L} \\
& \sum_{l \in \mathcal{L}} z_\ell \leq C && \\
& y^t_{\ell,b} \geq 0, x_\ell \in \mathbb{Z}_+, z_\ell \in \{0,1\}.
\end{aligned}
$$

The decision variables $y = \{y^t_{\ell,b} : t \in [T], \ell \in \mathcal{L}, b \in \mathcal{B}\}$ capture the matching decisions in this fluid setting, while the decision variables $x = \{x_\ell : \ell \in \mathcal{L}\}$ and $z = \{z_\ell : \ell \in \mathcal{L}\}$ encode the starting inventory levels and the set of stocked resources respectively. Working under arbitrary $W, C \in \mathbb{Z}_+$ for the remainder of this section, we use the triplet (y^*, x^*, z^*) to denote the optimal solution to Fluid-IP, and note that it is revenue management folklore (Talluri and Van Ryzin, 1998) that the objective value achieved by this optimal solution upper bounds the optimal expected reward, i.e. $\mathrm{Fluid}(W, C) \geq Z^*(W, C)$. The following lemma details how we can build a feasible solution to Fluid-IP that garners a constant fraction of the optimal objective value (condition (ii)) while ensuring that each customer type is matched with its ideal stocked resource, if matched at all (condition (i)).

Lemma 2.10. *From (y^*, x^*, z^*), one can construct a feasible solution $(\bar{y}, \bar{x}, \bar{z})$ to Fluid-IP such that*

(i) *We have that $\bar{y}^t_{\ell,b} > 0$ only if $b \in \mathcal{B}_\ell(\bar{x})$.*

(ii) $\displaystyle\sum_{t\in[T]} \sum_{\ell\in\mathcal{L}} \sum_{b\in\mathcal{B}} r_{\ell,b}\bar{y}^t_{\ell,b} \geq \frac{1}{2} \cdot \mathrm{Fluid}(W, C)$.

The proof of the above lemma very much follows the road map of how we select the starting inventory levels in Section 2.2; the starting inventory vector $\bar{x}$ is built by first considering the amount of fluid demand assigned to each resource under y^*, and then at most C of the least valuable units are potentially removed so as to ensure that the total number of units stocked is at most W.

The performance guarantee. The vector $\bar{x}$ dictates our choice for the starting inventory levels, which seeds the BoR matching policy. The following lemma gives the performance guarantee of this approach.

Lemma 2.11. $\hat{V}_1(\bar{x}) \geq \frac{1}{4} \cdot Z^*(W, C)$.

The guarantee stated above is established with a slight twist of the arguments used in Section 2.2, which is needed to make the balancing of the two lower bounds go through in a setting with general inventory constraints.

3 Computational Experiments

In this section, we conduct an extensive set of numerical experiments in which the approaches of Sections 2 are benchmarked against two formidable heuristics using a diverse array of randomly generated problem instances.

3.1 Computational Set-up

Instance generator. We randomly generate test instances with $L \in \{10, 25, 50\}$ resources and $B \in \{10, 25, 50\}$ customer types. We fix $T = 50$, and assume that there is exactly one arrival in each period. For every combination $(L, B) \in \{10, 25, 50\} \times \{10, 25, 50\}$, we generate 10 problem instances in total, which are each characterized by a distinct set of matching rewards and customer arrival probabilities that are generated as follows.

- *matching rewards:* The rewards are generated from a lognormal distribution with mean zero and scale parameter one.
- *arrival probabilities:* The arrival probabilities for each time period are uniformly generated from the m-dimensional probability simplex.

For each problem instance, we solve the inventory selection and online matching problem for $(W, C) \in \{10, 20, 50\} \times \{3, 5, 7\}$ using the approaches outlined next.

Implemented approaches. For each problem instances, and for each (W, C) pair under consideration, we implement the following three approaches, which are each labeled in relation to their corresponding approach for the online matching problem.

- Our approaches (**BoR**): With regard to the initial inventory decision, for the instances with $W = 50$, we use the approach of Section 2.2, since these instances fall within the sufficient supply regime. For all other instances, we choose the initial inventory vector using the algorithm outlined in Section 2.3. Next, for each instance, we implement an "inventory-adjusted" version of the BoR matching policy described in Section 2.1, in which the partitioning by ideal resource is carried out with respect to the current inventory level, rather than the initial inventory vector. This policy is formalized in Appendix D, where we show that this adjustment can only improve the performance of our original BoR matching policy, and hence all of the theoretical guarantees established in Section 2 continue to hold.
- Bid Price (**BID**): This heuristic is motivated by the classical bid-price heuristic often utilized for network revenue management (Talluri and Van Ryzin, 1998). Specifically, we solve Fluid-IP to determine the starting inventory level for each resource. Then, fixing these starting inventory levels, we re-solve Fluid-IP (now as a linear program), and store the dual variables $\{\mu_l\}_{\ell \in \mathcal{L}}$ associated with each of the resource constraints. Finally, we make our online matching decisions by replacing $V_{t+1}(X_t) - V_{t+1}(X_t - e_\ell)$ with μ_l within (1).
- Linear value function approximation (**LA**): For this approach, we again solve Fluid-IP to determine the starting inventory level for each resource, and then for the subsequent online matching problem, we implement the linear value function approximation of Ma et al. (2020),

which can be easily adapted to our matching setting. In short, this algorithm uses the linear value function approximation $V_t(X_t) \approx \sum_{\ell \in \mathcal{L}(X_t)} \mu_{\ell,t} x_{\ell,t}$, where $\mu_{\ell,t}$ are inventory-independent tuning parameters that represent an estimate of the value of a single unit of product ℓ in period t. The tuning parameters $\{\mu_{\ell,t}\}_{\ell \in \mathcal{L}, t \in [T]}$ are computed via a simple recursive expression. Similar to the bid price policy, we make our online matching decisions by replacing $V_{t+1}(X_t) - V_{t+1}(X_t - e_\ell)$ with $\mu_{\ell,t}$ within (1). As such, it is sensible to think of $\mu_{\ell,t}$ as a time-dependent bid-price.

3.2 Results.

The results of our experiments are presented in Tables 1a, 1b and 1c, which display the percent optimality gap, averaged across the 10 instances, for each of the three tested approaches. For ALG $\in \{\mathrm{BoR}, \mathrm{BID}, \mathrm{LA}\}$, we compute the optimality gap for each problem instance as $100 \cdot (\mathrm{Fluid}(W, C) - \mathbb{E}[\mathrm{ALG}])/\mathrm{Fluid}(W, C)$, where $\mathbb{E}[\mathrm{ALG}]$ is the expected reward of the approach as measured by 1,000 trials of Monte Carlo simulation. For every problem instance, each of the three approaches executed within a few minutes.

The results of these experiments reveal a handful of interesting insights and trends. However, before jumping into these trends, it is worth noting that optimality gaps in the range of 10% are also observed in the numerical experiments of Ma et al. (2020), and hence the magnitude of the optimality gaps reported in Tables 1a, 1b and 1c align with those typically observed in the literature. There are, however, certain parameter combinations (the $W = 10$ instances) that yield optimality gaps of over 20% across all three approaches. We attribute these large gaps to the difficulty of the problem when inventory is scarce, and also to the likelihood that $\mathrm{Fluid}(W, C)$ is a loose upper bound for these cases. Moving to the observed trends in our results, we see that our approach performs best uniformly across all combinations of W and C; outperforming BID by 10% and LA by 2% on average. We also observe that the optimality gaps across all three approaches generally grow as C is increased, indicating that the inherent difficulty of our joint inventory and online matching problem scales proportionally to C. Finally, for a fixed C, we see that the optimality gaps of all three approaches shrink significantly as we move from $W \in \{10, 20\}$ to the sufficient supply setting where $W = 50$. This trend likely follows from the fact that the online matching problem is more difficult when inventories are scarce.

B	W	C	Avg. % Opt. Gap BoR	BID	LA
10	10	3	9.34	23.46	10.24
10	10	5	9.34	23.46	10.24
10	10	7	9.34	23.46	10.24
10	20	3	8.15	26.71	9.65
10	20	5	9.03	27.17	10.57
10	20	7	9.03	27.17	10.57
10	50	3	3.68	6.73	6.71
10	50	5	6.57	9.45	7.76
10	50	7	8.03	10.55	8.31
25	10	3	12.92	23.41	15.14
25	10	5	14.28	23.95	15.91
25	10	7	14.18	24.06	15.81
25	20	3	8.09	26.36	10.84
25	20	5	11.65	27.05	14.28
25	20	7	12.13	27.16	14.83
25	50	3	3.99	6.71	7.71
25	50	5	6.53	9.50	8.17
25	50	7	8.69	11.16	9.24
50	10	3	13.78	25.49	16.63
50	10	5	18.69	26.23	20.13
50	10	7	21.87	26.58	20.67
50	20	3	8.39	27.61	11.56
50	20	5	12.56	28.57	15.51
50	20	7	15.10	28.97	17.61
50	50	3	3.97	6.63	8.05
50	50	5	6.55	9.45	8.67
50	50	7	8.63	11.44	9.66

(a) Instances with $L = 10$

B	W	C	Avg. % Opt. Gap BoR	BID	LA
10	10	3	10.39	26.83	11.19
10	10	5	10.39	26.83	11.19
10	10	7	10.39	26.83	11.19
10	20	3	9.06	27.43	10.07
10	20	5	10.39	28.65	11.28
10	20	7	10.39	28.65	11.28
10	50	3	4.48	7.82	8.07
10	50	5	7.53	10.95	8.67
10	50	7	9.27	13.06	9.55
25	10	3	14.07	25.77	15.68
25	10	5	18.25	28.22	18.47
25	10	7	17.87	28.69	18.21
25	20	3	8.78	28.56	10.79
25	20	5	12.56	29.44	14.55
25	20	7	15.83	30.57	16.78
25	50	3	3.60	7.33	7.99
25	50	5	6.25	10.58	8.69
25	50	7	8.81	12.91	10.07
50	10	3	13.98	24.93	16.46
50	10	5	18.97	16.74	20.14
50	10	7	21.11	27.97	22.00
50	20	3	8.01	27.04	10.93
50	20	5	12.45	28.33	17.59
50	20	7	15.98	28.98	17.59
50	50	3	3.49	6.46	7.74
50	50	5	6.53	9.80	8.51
50	50	7	8.66	11.95	9.62

(b) Instances with $L = 25$

B	W	C	Avg. % Opt. Gap BoR	BID	LA
10	10	3	10.43	25.50	11.21
10	10	5	10.43	25.50	11.21
10	10	7	10.43	25.50	11.21
10	20	3	9.36	28.15	10.69
10	20	5	11.15	29.35	11.75
10	20	7	11.06	29.35	11.75
10	50	3	3.56	7.46	6.54
10	50	5	7.40	11.36	8.43
10	50	7	9.77	13.56	10.40
25	10	3	15.32	27.93	16.79
25	10	5	18.40	30.02	19.51
25	10	7	17.91	31.64	18.91
25	20	3	8.95	28.00	11.10
25	20	5	13.64	30.31	15.47
25	20	7	16.15	31.64	17.49
25	50	3	3.69	7.13	7.93
25	50	5	6.36	10.56	8.54
25	50	7	8.66	13.05	9.65
50	10	3	13.82	27.24	16.13
50	10	5	20.36	29.41	21.79
50	10	7	23.92	31.38	24.17
50	20	3	8.24	29.27	10.81
50	20	5	12.60	30.18	14.88
50	20	7	16.65	32.15	18.34
50	50	3	3.33	6.80	8.28
50	50	5	6.00	10.36	8.78
50	50	7	8.28	13.06	10.11

(c) Instances with $L = 50$

Table 1: Average optimality gaps of the tested approaches.

4 The Anheuser Busch Inbev Trailer Problem

In this section, we first present an application of our problem setting as it relates to the logistical operations at ABI. To begin, we provide a high level overview of the ABI problem landscape, which is followed by a formal description of how the ABI Trailer problem can be viewed as a special case of the general problem framework considered in Section 2. Next, in Section 4.2, we describe how we tailor our approach for the sufficient supply setting to the ABI instances. Finally, in Sections 4.3 and 4.4, we present numerical experiment in which we then test our approach on realistic instances of the ABI Trailer Problem that are generated using real data from two brewery warehouses in Cartersville, Georgia (CRTV) and Fort Collins, Colorado (FCL).

4.1 Problem Landscape

ABI brews and packages its beer in multiple locations throughout the United States. After packaging, the finished product is transported to beer wholesalers via trucks provided by numerous third party logistics providers (3PLs). In an effort to reduce truck waiting times and increase transportation capacity, ABI prepares drop trailers, which are preloaded trailers of beer whose weights have been chosen in advance of the arrival of the third party trucks. Typically, drop trailers are preloaded 4hrs - 48hrs in advance of truck arrivals and comprise about 80% of the volume shipped from each brewery [3]. For each third party truck that arrives to their brewery warehouse, ABI must match this truck with a trailer of beer so as to maximize the shipping reward, which is calculated as a value that is proportional to the total volume of shipped beer. The maximum volume of beer that can be shipped by a truck is determined by the truck's weight and the federal law that the gross weight of a truck and trailer cannot exceed 80,000 lbs. If the gross weight of the assigned trailer and truck exceeds 80,000 lbs., then the truck must return to the loading dock for adjustment in order to comply with the federal law. This action is termed a "scaleback", and it causes additional costs for ABI since labor is required to remove the excess beer.

At the start of each day, ABI knows precisely the number of 3PL trucks scheduled to arrive to its warehouse, and in anticipation, prepares exactly this quantity of preloaded drop-trailers to be matched as trucks arrive. Although ABI may generally know how many trucks are arriving and have access to their arrival schedules, they typically do not have access to the exact weight of each truck until it arrives to the warehouse. The source of this uncertainty is two-fold. First, in

[3]The remaining 20% of volume is accounted for by "live trailers", which are loaded on the spot when a delivery truck arrives. These live trailers are generally used to accommodate arriving trucks that have non-standard set-ups for carrying the trailers.

the current environment, the trucking companies bid on many jobs simultaneously, and are thus constantly shuffling the type of trucks that they send to the various jobs they end up fulfilling. As a result, ABI never knows the exact weight of each arriving truck, only that it will be within some reasonable, but wide, range. On top of this, the trucks available to each carrier are dynamically changing throughout the day due to malfunctions and unpredictable traffic/weather patterns, which also adds to the difficulty of committing specific trucks to pick up specific loads in advance. Second, even if ABI knows the year, make, and model of a truck scheduled to arrive, the exact weight of the truck still possesses some level of variability due to the fact that (i) different materials may be used for some components of a truck [4], and (ii) additional equipment or fixtures are often added and removed from the trucks, since the 3PL companies service a wide range of customers. As such, relying only on historical truck arrival data, warehouse managers must first select the weight of each preloaded trailer, and then decide on a policy through which the trailers are matched to arriving trucks, with the goal across both sets of decisions being to maximize the total shipping reward across all truck arrivals. In what follows, we formalize the relationship between the drop trailer portion of ABI's shipping system and our general inventory selection and online matching problem.

The resources, customer types and matching rewards. In the ABI context, we replace the notion of a customer type by a truck type (i.e., a truck of a particular weight), and the notion of a resource by a trailer type (i.e., we refer to each set of trailers that is loaded at a unique weight as a trailer type). For trucks of type $b \in \mathcal{B}$, we let Ω_b denote their weight, and for trailers of type $\ell \in \mathcal{L}$, we let w_ℓ denote their weight. In Section 4.3, we discuss how we use historical data from ABI to derive the sets $\mathcal{B}$ and $\mathcal{L}$, as well the time-dependent arrival probabilities of each truck type. ABI receives a reward of $r_{\ell,b}$ for matching a trailer of type ℓ to a truck of type b, which takes the following form

$$r_{\ell,b} = \begin{cases} w_\ell \cdot r, & \text{if } w_\ell + \Omega_b \leq 80,000 \\ (80,000 - \Omega_b) \cdot r - ((w_\ell + \Omega_b) - 80,000) \cdot c & \text{otherwise.} \end{cases} \tag{7}$$

The matching rewards can be interpreted as follows. An arriving truck of type b has a total loading capacity of $80,000 - \Omega_b$ pounds, i.e., this is the maximum amount of beer that can be shipped by this particular truck type. The goal of ABI's logistical team is to fully utilize this available capacity, and hence ship as much beer as possible downstream to its wholesalers. To assess the efficacy of

[4] https://oversize.io/regulations/overweight-shipping-container-guide

any particular match, they estimate the per-pound value r of each pound loaded onto a trailer, as well as the per-pound cost c associated with a scaleback. As such, the term $r_{\ell,b}$ is computed as the total value in shipping $\min\{w_\ell, 80,000 - \Omega_b\}$ pounds of beer, minus any costs attributed to a scaleback event, which ensues if the particular matching results in a violation of the 80,000 lbs. weight limit. As discussed in Section 4.3, ABI has provided us with its estimates of both r and c for various warehouses and carriers.

Inventory considerations and ABI's current practice. We first note that ordering/production decisions for the ABI Trailer Problem are assumed to be exogenous, since ABI receives orders from wholesalers six weeks in advance of the actual delivery, and a large component of these orders consists of high demand products such as Bud Light, which are shipped to most wholesalers in extremely high volumes. Therefore, by the time that a production run is completed, there is always a large quantity of beer that needs to be shipped to the wholesalers. At this point, all inventory/production costs are considered sunk [5], and the sole operational focus is on selecting the trailer types to load, and deciding upon a policy to match these trailers to incoming third party trucks. Focusing on this two-stage problem, we assume that there are T scheduled arrivals of third party trucks, and hence we consider a setting with T discrete time periods, during which there is exactly one truck arrival in each period. As noted above, ABI does indeed generally know the number of arriving trucks each day, however in practice, the weight of each truck is only revealed once the truck arrives to the warehouse.

The current practice at ABI is to set $W = T$, meaning that ABI will always choose to preload exactly T trailers. Furthermore, ABI also currently loads only a single trailer type ($C = 1$), in which case the sole operational decision is to choose the weight of this lone trailer type [6], which will then be matched to all arriving trucks. Our approach provides the means to move beyond this simple setting, and to investigate the potential benefits of stocking multiple trailer types. That said, since ABI has never implemented a policy that loads more than a single trailer type, it is not clear how C should be selected in practice. For example, even if there are only trivial labor/resource-based costs associated with stocking additional trailer types, there still might be reason to select $C < L$. Namely, perhaps the logistics of stocking L distinct trailer types, and then properly managing their respective inventories throughout the matching process, is too cumbersome, however, it is

[5] ABI keeps at most two days of inventory on-hand in each warehouse, and hence holding costs are a trivial consideration.

[6] This problem can be reduced to a standard newsvendor problem.

quite manageable to handle 5 different trailer types. In this case, ABI should seek the best initial inventory vector that uses at most 5 trailer types, which is precisely the solution that our constrained framework yields. Alternatively, ABI may know that there will be substantial operational costs associated with stocking multiple trailer types, but since they have never considered implementing such a policy, they do not have precise estimates of these costs. To check whether it is indeed worthwhile to estimate these costs, ABI needs to understanding the extent to which matching rewards can be improved as C is increased. In this case, our constrained approach can be carried out for various values of C to assess the marginal improvement that can result as ABI considers utilizing more and more distinct trailer types.

4.2 A Practical Approach for the ABI Trailer Problem

Before detailing the approach we implement for the ABI-specific instances, we first show an intriguing structural result related to the optimal matching policy in this setting. Specifically, due to the piecewise linear structure of the matching rewards, we are able to show that the optimal matching decision in each period, and for each truck, can always be reduced to a choice between two specific trailers.

The updated optimal matching policy. To begin, we update the dynamic program in (1) to reflect the fact that, for the ABI setting, arriving trucks cannot be rejected. With this update, the optimal matching policy can be derived via the following dynamic program

$$V_t(X_t) \quad = \quad \sum_{b \in \mathcal{B}} p_{b,t} \cdot \max_{\ell \in \mathcal{L}(X_t)} \{r_{\ell,b} + V_{t+1}(X_t - e_\ell)\} + \left(1 - \sum_{b \in \mathcal{B}} p_{b,t}\right) \cdot V_{t+1}(X_t), \tag{8}$$

whose value functions mirror those of (1), with the update that each arriving truck must be assigned an available resource. It is important to note that we are unaware of any results that establish performance guarantees for online matching problems where the matching rewards depend on both the customer and matched resource, and when rejections are not permitted. To further cement the difficulty of this no-reject setting, in Appendix C.2, we show that the general version of our joint inventory and online matching problem in which arriving customers cannot be rejected is NP-Hard to approximate with any constant factor $\alpha > 0$.

The greedy pick-two policy. In what follows, we formalize the structural result concerning the optimal matching policy mentioned above, and further discuss its implications. To begin, for arbitrary inventory vector $X \in \mathbb{Z}_+^L$ and truck $b \in \mathcal{B}$, define

$$\ell^{\uparrow}(b, X) = \underset{\substack{\ell \in \mathcal{L}(X): \\ w_\ell + \omega_b \geq 80{,}000}}{\operatorname{argmin}} \quad w_\ell$$

and

$$\ell^{\downarrow}(b, X) = \underset{\substack{\ell \in \mathcal{L}(X): \\ w_\ell + \omega_b \leq 80{,}000}}{\operatorname{argmax}} \quad w_\ell.$$

In other words, $\ell^{\uparrow}(b, X)$ is the minimum weight trailer that, if assigned to truck b, induces a scaleback, while $\ell^{\downarrow}(b, X)$ is the maximum weight trailer that does not induce a scaleback if assigned to truck b. Note that we can introduce dummy trailer types indexed 0 and $L+1$ with infinite capacities and respective weight of $-\infty$ and ∞ so that $\ell^{\uparrow}(b, X_t)$ or $\ell^{\downarrow}(b, X_t)$ are well-defined. The following lemma, whose proof is presented in Appendix B, reveals that for any period $t \in [T]$, inventory vector X_t, and arriving truck $b \in \mathcal{B}$, the optimal matching policy must select either $\ell^{\uparrow}(b, X_t)$ or $\ell^{\downarrow}(b, X_t)$.

Lemma 4.1. *For any period $t \in [T]$, inventory vector X_t, and arriving truck $b \in \mathcal{B}$, we have that*

$$\max_{\ell \in \mathcal{L}(X_t)} \{r_{\ell,b} + V_{t+1}(X_t - e_\ell)\} = \max_{\ell \in \{\ell^{\uparrow}(b, X_t), \ell^{\downarrow}(b, X_t)\}} \{r_{\ell,b} + V_{t+1}(X_t - e_\ell)\},$$

where the value functions $V_t(\cdot)$ are defined as in (8)*.*

It is important to note that, while Lemma 4.1 narrows down the set of candidate optimal trailers from $|\mathcal{L}(X_t)|$ to 2, optimally choosing between $\ell^{\uparrow}(b, X_t)$ and $\ell^{\downarrow}(b, X_t)$ in each period is still a nontrivial task, as doing so requires oracle access to the value functions defined in (8). Nonetheless, a natural heuristic approach is to select between these two candidates by merely choosing the one that garners the higher immediate reward. We refer to this policy as the greedy pick-two policy. We explain below that such a greedy policy is also closely related to our inventory-adjusted BoR matching policy introduced in the general problem setting, hence it is a good candidate policy to test for the ABI Trailer Problem.

The implemented approach. Since the ABI instances fall within the framework of the sufficient supply setting, we choose initial inventory levels of the trailers as described in Section 2.2, albeit with the following updated version of Claim 2.7, which reveals an improved (in relation to X'_C) set of trailers to stock.

Claim 4.2. *For the ABI Trailer Problem, there exists a polynomial-time algorithm that constructs an initial inventory vector* $X_C^* \in \mathcal{F}(\infty, C)$ *that satisfies*

$$\hat{V}_1(X_C^*) = Z^*(\infty, C).$$

With Claim 4.2, we can derive $\bar{X}_1$ using (6), where X_C' is replaced with X_C^*. We then adopt the inventory-adjusted BoR matching policy of Appendix D where the reject option is removed. It is easy to see that with this adjustment, the inventory-adjusted BoR matching policy simply matches each arriving truck with the highest-reward trailer type, among those that are available. More formally, if the current inventory level in period t is X_t, then we match an arriving truck of type $b \in \mathcal{B}$ to resource $\ell(b, X_t)$, i.e. the ideal trailer type from among those that are currently in-stock. Since we do slightly deviate from the inventory-adjusted BoR matching policy, the performance guarantee of Theorem 2.4 no longer holds in general. Specifically, with this slight adjustment to the BoR policy, it is not hard to see that the corresponding value functions associated with this policy no longer satisfy the structural properties presented in Claim 2.2, which are critical for establishing the two lower bounds presented in Section 2.2. That said, Lemma 4.1 provides ample theoretical support to adopt such a policy. Specifically, given the structure of the piecewise linear matching rewards, it is fairly straightforward to see that $\ell(b, X_t) = \mathrm{argmax}_{\ell \in \{\ell^{\uparrow}(b,X_t), \ell^{\downarrow}(b,X_t)\}} r_{\ell,b}$, and hence this myopic policy reduces precisely to the greedy pick-two policy defined above.

4.3 Experimental Set-up

In this section, we provide a detailed overview of the data set that has been provided to us by ABI, which guides our parameter selection and gives a sense of the scale of the problems that we consider. Our experiments investigate the inventory selection and allocation problem for ABI on a single day, during which there are many arriving trucks that need to be matched to pre-loaded trailers of beers (e.g., resources).

ABI Data Description. For each of the two warehouses at Cartersville, Georgia (CRTV) and Fort Collins, Colorado (FCL), we have access to historical third party truck arrival data from February to July of 2016 that allows us to simulate realistic instances of the ABI Trailer Problem. For each truck arrival, we have a timestamp giving the truck's arrival time and date, the weight of the truck and its associated carrier, which denotes the unique third party delivery service to which it belongs. Since ABI typically preloads a distinct set of trailers for each carrier at the beginning of the day, we solve separate ABI Trailer Problems for each warehouse-carrier pair. For each truck type $b \in \mathcal{B}$ (e.g., defined by a truck weight Ω_b), we set its time-dependent arrival probability $p_{b,t}$ to be the fraction of time that the t-th arrival of the day is of a type b truck. The number of arrivals T is assumed to be the maximum number of arrivals that are observed on any single day over the six months of historical arrivals. To ensure that we consider the more complex problem instances, we ignore carriers that always had fewer than 20 arrivals on any given day. Finally, estimates of the per-pound value r and the scaleback cost c were provided to us by ABI for each carrier. A full summary of these parameters for each carrier at each warehouse is given in Tables 2a and 2b.

carrier	T	c	r	min Ω_b	max Ω_b	mean Ω_b	stdev. Ω_b
GTGA	23	0.047	0.007	16,420	20,220	18,106.27	890.31
MTNF	22	0.047	0.008	15,820	19,480	16,987.87	631.29
PRIJ	33	0.047	0.015	17,420	21,160	19,172.68	915.89
WENX	34	0.047	0.012	14,880	20,280	17,427.38	1,765.57

(a) CRTV

carrier	T	c	r	min Ω_b	max Ω_b	mean Ω_b	stdev. Ω_b
TAMI	25	0.050	0.005	17,860	21,380	19,545.87	678.77
WENP	36	0.050	0.008	15,160	21,860	19,412.49	1,145.80
WERD	38	0.050	0.004	14,000	22,460	16,727.14	1,759.37
WERS	24	0.050	0.005	13,500	20,000	16,390.20	1,334.95

(b) FCL

Table 2: Parameters of the ABI Trailer Problem at warehouses CRTV and FCL.

Final implementation details. For each carrier at each warehouse, we set $\mathcal{L}$ to be trailers loaded at weights represented by a linearly spaced set of grid points in increments of 100 over the interval $[80K - \max_{b\in\mathcal{B}} \Omega_b, 80K - \min_{b\in\mathcal{B}} \Omega_b]$. It is easy to see that this range of potential trailer types includes the smallest and largest trailer weights that one would ever consider choosing. This linearly spaced set of potential trailer weights leads to values of $L = |\mathcal{L}|$ that fall between 35 and 90. For each carrier at each of the two warehouses that we consider, and for each $C \in [L]$, we carry out the approach of Section 4.2 to choose the initial inventory in $\mathcal{F}(T, C)$, and then apply the greedy pick-two policy for matching. We measure the performance of our approach relative to $Z^*(\infty, C)$, which can be computed by solving the (Fluid-IP) with $W = LT$. Specifically, we report the optimality gap of our approach by $100 \times (Z^*(\infty, C) - \mathbb{E}[\text{ALG}])/Z^*(\infty, C)$, where $\mathbb{E}[\text{ALG}]$ is the expected reward of our approach estimated by 10,000 trials of Monte Carlo Simulations. All experiments used Python 3.6 on an Intel Core i5 with 3.2 GHz CPU and 32GB of RAM and Gurobi 6.5.1 as the integer programming solver.

4.4 Results

We report the performance of our approach for only $C \in \{1, 3, 5\}$ in Tables 3a and 3b, since we observed little marginal change in both the expected rewards and optimality gaps when C was increased beyond 5. In these two tables, column one gives the carrier, while column two specifies the value of C tested within the implementation of our algorithm. Column three gives the expected reward derived from our approach, which we estimate using 10,000 Monte Carlo simulations. Finally, column four shows the optimality gap of our approach in relation to $Z^*(\infty, C)$, where for $C = 1$, we report this gap as "NA" since our algorithm will always achieve an expected reward of $Z^*(\infty, 1)$ for these instances.

Carrier (L)	C	Exp. Rew.	% OPT GAP
GTGA (39)	1	9,876	NA
	3	9.938	0.63
	5	9,950	0.76
MTNF (38)	1	10,678	NA
	3	10,727	0.24
	5	10,736	0.28
PRIJ (39)	1	29,076	NA
	3	29,262	0.14
	5	29,278	0.29
WENX (55)	1	25,132	NA
	3	25,222	0.53
	5	25,211	0.81

(a) CRTV

Carrier (L)	C	Exp. Rew.	% OPT GAP
TAMI (37)	1	6,920	NA
	3	6,956	0.56
	5	6,947	0.88
WENP (68)	1	16,678	NA
	3	16,767	0.29
	5	16,799	0.29
WERD (86)	1	8,581	NA
	3	8,688	0.52
	5	8,690	0.84
WERS (66)	1	7,061	NA
	3	7,108	0.97
	5	7,093	1.45

(b) FCL

Table 3: Performance metrics of our approach at warehouses CRTV and FCL.

We would like to highlight a few intriguing observations from these results. First, we observe that our approach is near optimal for almost all test cases; for $C = 3$ and $C = 5$, the average optimality gap at warehouse CRTV is 0.38% and 0.54%, and the average gap at FCL is a measly 0.58% and 0.86%. Furthermore, even when we consider the full spectrum of instances tested beyond those reported in Tables 3a and 3b (i.e., we consider $C \in [L]$), we observe a worst case optimality gap of 2.64% for warehouse FCL, carrier WERS, and $C = 21$. Finally, we note that for a fixed warehouse and carrier, it should become more difficult to compete against $Z^*(\infty, C)$ as C is increased, since the problem becomes more difficult as one does so. As such, it is no surprise that we see slight upticks in the percentage optimality gaps as C is increased from 1 to 5.

Next, we investigate the potential reward gains that can result by utilizing more than a single trailer type, which is the current practice at ABI. Figure 1 illustrates precisely this improvement for each of the eight carriers as C is increased up to 10. At each value of C reported on the x-axis of Figure 1, we report the maximum expected reward of our adjusted greedy policy (computed via Monte Carlo simulation) for an inventory vector that uses at most C types of resources. More specifically, for a particular value of C, we report the best expected revenue earned across all $C' \leq C$ tested, and hence the line plots in Figure 1 are guaranteed to be increasing in C. It is clear from Figure 1 that there is little marginal gain in increasing C beyond 5. Ultimately, we see that by stocking trailers of up to five distinct weights, ABI has the potential to improve upon its current practice by 0.35-1.29%. In fact, this 0.35-1.29% increase equates approximately to a \$200 increase per day for each carrier. Given that ABI has hundreds of carriers that service 21 breweries all over the U.S., this slight improvement in performance can improve logistical operations at ABI by millions of dollars per year.

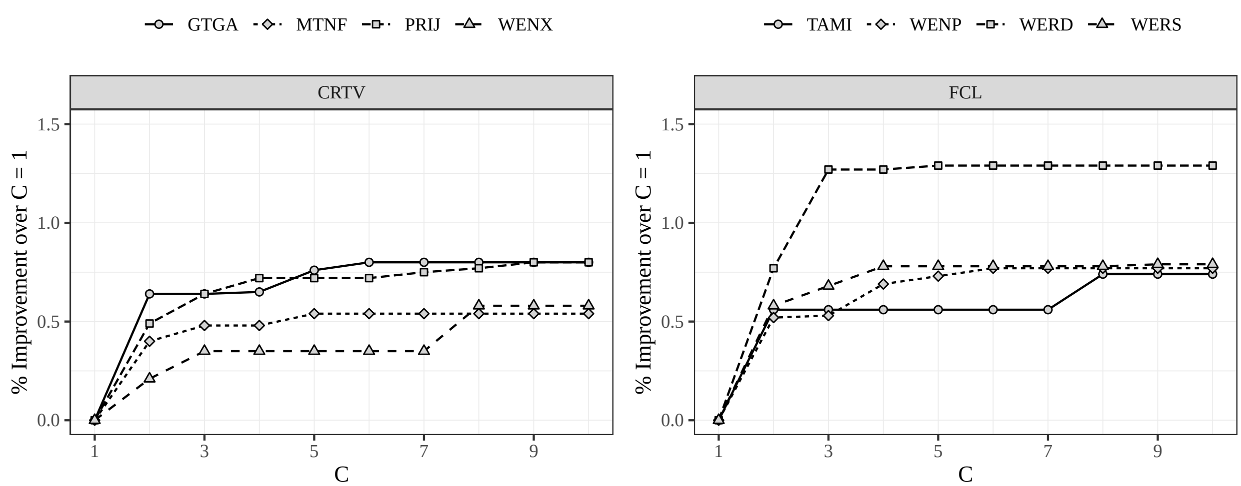

Figure 1: The percentage improvement in expected reward as C is increased beyond 1.

Incorporating Operational Costs. Our constrained formulation of the ABI Trailer Problem, coupled with the simple, efficient, and near-optimal solution approach we develop, can be employed to solve the following costed version of our problem, in which there are costs associated with the number of trailer types utilized and the total number of trailers loaded:

$$\max_{X_1 \in \mathbb{Z}_+^L} V_1(X_1) - \text{Cost}(\sum_{\ell \in \mathcal{L}} x_{\ell,1}, |\mathcal{L}(X_1)|). \tag{9}$$

In the above formulation, $\text{Cost}(\cdot, \cdot)$ captures the operational costs associated with choosing X_1, which are assumed to be an arbitrary function of (i) the total number of units stocked and (ii) the number of unique resources stocked. Since ABI has always stocked T units of single trailer type, they were not able to provide us with estimates of these costs. Nonetheless, if the cost function is known, we can solve this costed version of the problem via the following alternative formulation

$$\max_{\substack{W \in \{1,\ldots,LT\} \\ C \in \{1,\ldots,L\}}} \left\{ \max_{X_1 \in \mathcal{F}(W,C)} V_1(X_1) \right\} - \mathrm{Cost}(W, C). \tag{10}$$

It is easy to see that problems (9) and (10) are equivalent, i.e. they have the same optimal objective value and an optimal solution to one, can be easily converted to an optimal solution to the other. Furthermore, the outer maximization can be solved by enumerating over all L^2T possibilities for W and C, and the inner maximization is exactly problem (2). In Appendix E, we present extensive ABI-based numerical experiments in which we investigate the extent to which our solution approach for the constrained version of the problem can be exploited to solve (10) along the lines just-mentioned. We find that our approach continues to perform near-optimally, as we never observe optimality gaps exceeding 1.5%.

5 Conclusion

In this paper, we highlight the importance of jointly considering the initial inventory selection problem and the subsequent online resource allocation problem. As such, we hope that our work inspires future papers that tackle these two important operational problems jointly as well. Along this line, there are a number of directions for future work. First, it is interesting to ask if our performance guarantees for the general matching setting could be improved by using more sophisticated tools for the online matching problem. It is also intriguing to wonder if these sorts of approximation guarantees can be extended to a choice setting, where the DM selects an assortment of products to make available for consumption, and arriving customers choose amongst these products according to some pre-specified choice model.

References

Aouad, A., Levi, R. and Segev, D. (2018), 'Greedy-like algorithms for dynamic assortment planning under multinomial logit preferences', *Operations Research* **66**(5), 1321–1345.

Aouad, A. and Segev, D. (2019), 'The stability of mnl-based demand under dynamic customer substitution and its algorithmic implications', *Available at SSRN 3498325* .

Ball, M. O. and Queyranne, M. (2009), ‘Toward robust revenue management: Competitive analysis of online booking’, *Operations Research* **57**(4), 950–963.

Chan, C. W. and Farias, V. F. (2009), ‘Stochastic depletion problems: Effective myopic policies for a class of dynamic optimization problems’, *Mathematics of Operations Research* **34**(2), 333–350.

Cornuejols, G., Fisher, M. L. and Nemhauser, G. L. (1977), ‘Exceptional paper—location of bank accounts to optimize float: An analytic study of exact and approximate algorithms’, *Management science* **23**(8), 789–810.

DeValve, L., Pekeč, S. and Wei, Y. (2020), ‘A primal-dual approach to analyzing ato systems’, *Management Science* **66**(11), 5389–5407.

Dong, L., Kouvelis, P. and Tian, Z. (2009), ‘Dynamic pricing and inventory control of substitute products’, *Manufacturing & Service Operations Management* **11**(2), 317–339.

El Housni, O., Goyal, V., Humair, S., Mouchtaki, O., Sadighian, A. and Wu, J. (2021), ‘Joint assortment and inventory planning for heavy tailed demand’, *Available at SSRN 3832909* .

Feige, U. (1998), ‘A threshold of ln n for approximating set cover’, *Journal of the ACM (JACM)* **45**(4), 634–652.

Feldman, J. B. and Topaloglu, H. (2017), ‘Revenue management under the markov chain choice model’, *Operations Research* **65**(5), 1322–1342.

Feldman, J., Korula, N., Mirrokni, V., Muthukrishnan, S. and Pal, M. (2009), ‘Online ad assignment with free disposal’, *WINE* pp. 374–385.

G. L. Nemhauser, L. A. W. and Fisher, M. L. (1978), ‘An analysis of approximations for maximizing submodular set functions’, *Mathematical Programming* **14**, 265–294.

Gallego, G., Iyengar, G., Phillips, R. and Dubey, A. (2004), ‘Managing flexible products on a network’.

Gallego, G. and Kim, S. (2020), ‘Joint pricing and inventory decisions for substitutable products’, *Available at SSRN 3513274* .

Gallego, G., Ratliff, R. and Shebalov, S. (2015), ‘A general attraction model and sales-based linear program for network revenue management under customer choice’, *Operations Research* **63**(1), 212–232.

Gong, X.-Y., Goyal, V., Iyengar, G. N., Simchi-Levi, D., Udwani, R. and Wang, S. (2021), 'Online assortment optimization with reusable resources', *Management Science* .

Goyal, V., Levi, R. and Segev, D. (2016), 'Near-optimal algorithms for the assortment planning problem under dynamic substitution and stochastic demand', *Operations Research* **64**, 219–235.

Honhon, D., Gaur, V. and Seshadri, S. (2010), 'Assortment planning and inventory decisions under stockout-based substitution', *Operations Research* **58**, 1364–1379.

Jasin, S. and Kumar, S. (2012), 'A re-solving heuristic with bounded revenue loss for network revenue management with customer choice', *Mathematics of Operations Research* **37**(2), 313–345.

Karp, R. M. (1972), Reducibility among combinatorial problems, *in* 'Complexity of computer computations', Springer, pp. 85–103.

Liu, Q. and van Ryzin, G. (2009), 'On the choice based linear programming model for network revenue management', *Manufacturing Service Operations Management* **10**, 288–310.

Ma, Y., Rusmevichientong, P., Sumida, M. and Topaloglu, H. (2020), 'An approximation algorithm for network revenue management under nonstationary arrivals', *Operations Research* **68**(3), 834–855.

Mahajan, S. and Van Ryzin, G. (2001), 'Stocking retail assortments under dynamic consumer substitution', *Operations Research* **49**(3), 334–351.

Martínez-de Albéniz, V. and Kunnumkal, S. (2021), 'A model for integrated inventory and assortment planning', *Management Science* .

Mehta, A. (2013), 'Online matching and ad allocation', *Foundations and Trends in Theoretical Computer Science* **8**, 265–368.

Méndez-Díaz, I., Miranda-Bront, J., Vulcano, G. and Zabala, P. (2010), 'A branch-and-cut algorithm for the latent-class logit assortment problem', *Discrete Applied Mathematics* **36**, 383–390.

Previgliano, F. and Vulcano, G. (2021), 'Managing uncertain capacities for network revenue optimization', *Manufacturing & Service Operations Management* .

Reiman, M. I. and Wang, Q. (2015), ‘Asymptotically optimal inventory control for assemble-to-order systems with identical lead times’, *Operations Research* **63**(3), 716–732.

Saghafian, S., Trichakis, N., Zhu, R. and Shih, H. (2019), ‘Joint patient selection and scheduling under no-shows: Theory and application in proton therapy’, *Available at SSRN 3396096* .

Simpson, R. (1989), *Using Network Flow Techniques to Find Shadow Prices for Market Demands and Seat Inventory Control*, FTL memorandum, MIT, Department of Aeronautics and Astronautics, Flight Transportation Laboratory.
URL: *https://books.google.com/books?id=GUuvHAAACAAJ*

Stein, C., Truong, V.-A. and Wang, X. (2020), ‘Advance service reservations with heterogeneous customers’, *Management Science* **66**(7), 2929–2950.

Talluri, K. and Ryzin, G. V. (1999), ‘A randomized linear programming method for computing network bid prices’, *Transportation Science* **33**, 207–216.

Talluri, K. and Van Ryzin, G. (1998), ‘An analysis of bid-price controls for network revenue management’, *Management science* **44**(11-part-1), 1577–1593.

Talluri, K. and van Ryzin, G. (2004), ‘Revenue management under a general discrete choice model of consumer behavior’, *Management Science* **50**(1), 15–33.

Topaloglu, H. (2009), ‘Using lagrangian relaxation to compute capacity-dependent bid prices in network revenue management’, *Operations Research* **57**, 637–649.

Topaloglu, H. (2013), ‘Joint stocking and product offer decisions under the multinomial logit model’, *Production and Operations Management* **22**(5), 1182–1199.

Wang, X., Truong, V.-A. and Bank, D. (2018), ‘Online advance admission scheduling for services with customer preferences’, *arXiv preprint arXiv:1805.10412* .

Zhang, D. and Cooper, W. L. (2005), ‘Revenue management for parallel flights with customer-choice behavior’, *Operations Research* **53**(3), 415–431.

A Proofs from Section 2

A.1 Proof of Claim 2.1

We prove the result via induction over t. The base case of $t = T + 1$ based on the terminal cases of (3) and (4). For the general case of $t \in [T]$, we have that

$$\begin{aligned}
\hat{V}_t(X_t; X_1) &= \sum_{\ell \in \mathcal{L}(X_t)} \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot \left[r_{\ell,b} - (\hat{V}_{t+1}(X_t; X_1) - \hat{V}_{t+1}(X_t - e_\ell; X_1)) \right]^+ + \hat{V}_{t+1}(X_t; X_1) \\
&= \sum_{\ell \in \mathcal{L}(X_t)} \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot \left[r_{\ell,b} - (\hat{V}^\ell_{t+1}(x_{\ell,t}; X_1) - \hat{V}^\ell_{t+1}(x_{\ell,t} - 1; X_1)) \right]^+ + \sum_{\ell \in \mathcal{L}(X_t)} \hat{V}^\ell_{t+1}(x_{\ell,t}; X_1) \\
&= \sum_{\ell \in \mathcal{L}(X_t)} \hat{V}^\ell_t(x_{\ell,t}; X_1),
\end{aligned}$$

where the second inequality follows by the induction hypothesis.

A.2 Proof of Claim 2.2

The three properties of of Claim 2.2 regarding the decoupled value functions are restated below.

(i) $\hat{V}^\ell_t(x + 1; X_1) \geq \hat{V}^\ell_t(x; X_1)$ for inventory levels $x \geq 0$,

(ii) $\Delta\hat{V}^\ell_t(x; X_1) \geq \Delta\hat{V}^\ell_t(x + 1; X_1)$ for inventory levels $x \geq 1$,

(iii) $\Delta\hat{V}^\ell_t(x; X_1) \geq \Delta\hat{V}^\ell_{t+1}(x; X_1)$ for inventory levels $x \geq 1$.

We prove each of the results inductively over the time periods. First, note that each of the three properties holds trivially for time period $T + 1$. We begin with the proof of property (ii), which is the most involved.

Proof of property (ii): Using the value functions presented in (4), we have that

$$\begin{aligned}
\Delta\hat{V}^\ell_t(x; X_1) - \Delta\hat{V}^\ell_t(x + 1; X_1) &= \Delta\hat{V}^\ell_{t+1}(x; X_1) - \Delta\hat{V}^\ell_{t+1}(x + 1; X_1) \\
&\quad + \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot \left[r_{\ell,b} - \Delta\hat{V}^\ell_{t+1}(x; X_1) \right]^+ - \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot \left[r_{\ell,b} - \Delta\hat{V}^\ell_{t+1}(x - 1; X_1) \right]^+ \\
&\quad - \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot \left[r_{\ell,b} - \Delta\hat{V}^\ell_{t+1}(x + 1; X_1) \right]^+ + \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot \left[r_{\ell,b} - \Delta\hat{V}^\ell_{t+1}(x; X_1) \right]^+ . \\
&\geq \Delta\hat{V}^\ell_{t+1}(x; X_1) - \Delta\hat{V}^\ell_{t+1}(x + 1; X_1) \\
&\quad + \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot y^1_b \cdot \left(r_{\ell,b} - \Delta\hat{V}^\ell_{t+1}(x; X_1) \right) - \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot y^1_b \cdot \left(r_{\ell,b} - \Delta\hat{V}^\ell_{t+1}(x - 1; X_1) \right) \\
&\quad - \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot y^2_b \cdot \left(r_{\ell,b} - \Delta\hat{V}^\ell_{t+1}(x + 1; X_1) \right) + \sum_{b \in \mathcal{B}_\ell(X_1)} p_{b,t} \cdot y^2_b \cdot \left(r_{\ell,b} - \Delta\hat{V}^\ell_{t+1}(x; X_1) \right)
\end{aligned}$$

where we define

$$y_b^1 = \begin{cases} 1 & \text{if } r_{\ell,b} > \Delta V_{t+1}^l(X_1, x-1) \\ 0 & \text{otherwise,} \end{cases}$$

and

$$y_b^2 = \begin{cases} 1 & \text{if } r_{\ell,b} > \Delta V_{t+1}^l(X_1, x+1) \\ 0 & \text{otherwise.} \end{cases}$$

Note that the inequality above follows because y_b^1 and y_b^2 are feasible policies for when the current inventory levels is x. After some algebra and cancelling and grouping of common terms, we get

$$\begin{aligned}
\Delta\hat{V}_t^\ell(x;X_1) - \Delta\hat{V}_t^\ell(x+1;X_1) \geq\ & \Delta\hat{V}_{t+1}^\ell(x;X_1) - \Delta\hat{V}_{t+1}^\ell(x+1;X_1) \\
& - \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot y_b^1 \cdot \Big(\Delta\hat{V}_{t+1}^\ell(x;X_1) - \Delta\hat{V}_{t+1}^\ell(x-1;X_1)\Big) \\
& + \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot y_b^2 \cdot \Big(\Delta\hat{V}_{t+1}^\ell(x+1;X_1) - \Delta\hat{V}_{t+1}^\ell(x;X_1)\Big) \\
& = (1 - \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot y_b^2)\cdot \Big(\Delta\hat{V}_{t+1}^\ell(x;X_1) - \Delta\hat{V}_{t+1}^\ell(x+1;X_1)\Big) \\
& + \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot y_b^2 \cdot \Big(\Delta\hat{V}_{t+1}^\ell(x-1;X_1) - \Delta\hat{V}_{t+1}^\ell(x;X_1)\Big) \\
& \geq 0,
\end{aligned}$$

where the last inequality follows by the induction hypothesis.

Proof of property (i): Using the value functions presented in (4), we observe that

$$\begin{aligned}
\hat{V}_t^\ell(x+1;X_1) - \hat{V}_t^\ell(x;X_1) =& \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot \Big[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x+1;X_1)\Big]^+ - \\
& \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot \Big[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x;X_1)\Big]^+ + \Big(\hat{V}_{t+1}^l(x+1;X_1) - \hat{V}_{t+1}^l(x;X_1)\Big) \\
& \geq \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot \Big[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x+1;X_1)\Big]^+ - \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot \Big[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x;X_1)\Big]^+ \\
& \geq \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot \Big[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x;X_1)\Big]^+ - \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot \Big[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x;X_1)\Big]^+ \\
& = 0.
\end{aligned}$$

The first inequality uses the induction hypothesis and the second uses property (ii).

Proof of property (iii): Again using the value functions presented in (4), we observe that

$$\begin{aligned}
\Delta\hat{V}_t^\ell(x;X_1) =& \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot \Big[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x-1;X_1)\Big]^+ - \\
& \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot \Big[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x-2;X_1)\Big]^+ + \Delta\hat{V}_{t+1}^\ell(x;X_1)
\end{aligned}$$

and so

$$\begin{aligned}
\Delta\hat{V}_t^\ell(x;X_1) - \Delta\hat{V}_{t+1}^\ell(x;X_1) = &\sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\left[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x-1;X_1)\right]^+ - \\
&\sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\left[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x-2;X_1)\right]^+ \\
&\geq \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\left[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x-2;X_1)\right]^+ - \\
&\sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\left[r_{\ell,b} - \Delta\hat{V}_{t+1}^\ell(x-2;X_1)\right]^+ \\
&= 0.
\end{aligned}$$

The lone inequality uses property (iii) of the claim.

A.3 Proof of Claim 2.3

To begin, note that version of our simple matching policy for starting inventory vector X_1 and resource $\ell\in\mathcal{L}(X_1)$ can be expressed through binary decision variables $y_{b,t}^\ell$, which take value 1 if customer $b\in\mathcal{B}_\ell(X_1)$ is assigned to resource ℓ during time period t. Under the optimal policy, we have that

$$y_{b,t}^\ell = \begin{cases} 1 & \text{if } \; r_{\ell,b} > \Delta V_{t+1}^\ell(x;X_1) \\ 0 & \text{otherwise,} \end{cases}$$

which can easily be teased out from the value functions presented in (4). Moreover for starting inventory vector X_1', we can construct a feasible matching policy $\hat{y}_{b,t}^\ell$ for each $b\in\mathcal{B}_l(X_1')$ and $l\in\mathcal{L}(X_1')$ as follows

$$\hat{y}_{b,t}^\ell = \begin{cases} y_{b,t}^\ell & \text{if } \; l\in\mathcal{L}(X_1')\cap\mathcal{L}(X_1) \\ 0 & \text{otherwise,} \end{cases}$$

which is a well defined simple matching policy since $\mathcal{B}_\ell(X_1)\subseteq\mathcal{B}_l(X_1')$. Next, we prove the desired result by induction over the time periods. The base case holds trivially for period $T+1$. For time period $t\in[T]$, we observe that for any $l\in\mathcal{L}(X_1')\cap\mathcal{L}(X_1)$ and inventory level $x\in\mathbb{Z}_+$, we have that

$$\begin{aligned}
\hat{V}_t^\ell(x;X_1) &= \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot y_{b,t}^\ell\cdot\left(r_{\ell,b} + \hat{V}_{t+1}^\ell(x-1;X_1)\right) + \left(1 - \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot y_{b,t}^\ell\right)\hat{V}_{t+1}^\ell(x;X_1) \\
&= \sum_{b\in\mathcal{B}_l(X_1')} p_{b,t}\cdot \hat{y}_{b,t}^\ell\cdot\left(r_{\ell,b} + \hat{V}_{t+1}^\ell(x-1;X_1)\right) + \left(1 - \sum_{b\in\mathcal{B}_l(X_1')} p_{b,t}\cdot \hat{y}_{b,t}^\ell\right)\hat{V}_{t+1}^\ell(x;X_1) \\
&\leq \sum_{b\in\mathcal{B}_l(X_1')} p_{b,t}\cdot \hat{y}_{b,t}^\ell\cdot\left(r_{\ell,b} + \hat{V}_{t+1}^\ell(x-1;X_1')\right) + \left(1 - \sum_{b\in\mathcal{B}_l(X_1')} p_{b,t}\cdot \hat{y}_{b,t}^\ell\right)\hat{V}_{t+1}^\ell(x;X_1') \\
&\leq \hat{V}_t^\ell(X_1',x).
\end{aligned}$$

The second equality results since the policy $\hat{y}_{b,t}^{\ell}$ mirrors the policy $y_{b,t}^{\ell}$. The first inequality results by the induction hypothesis and the second due the fact that the policy $\hat{y}_{b,t}^{\ell}$ is feasible but not necessarily optimal.

A.4 Proof of Lemma 2.5

We have that

$$\begin{aligned}
\hat{V}_1(X_1) &= \sum_{\ell\in\mathcal{L}(X_1)} \hat{V}_1^{\ell}(x_{\ell,1};X_1) \\
&= \sum_{\ell\in\mathcal{L}(X_1)} \sum_{x=1}^{x_{\ell,1}} \left(\hat{V}_1^{\ell}(x;X_1) - \hat{V}_1^{\ell}(x-1;X_1)\right) \\
&= \sum_{\ell\in\mathcal{L}(X_1)} \sum_{x=1}^{x_{\ell,1}} \Delta\hat{V}_1^{l}(x;X_1) \\
&\geq \sum_{\ell\in\mathcal{L}(X_1)} x_{\ell,1}\cdot\Delta\hat{V}_1^{\ell}(x_{\ell,1};X_1).
\end{aligned}$$

where the inequality follows since properties 1 and 2 of Claim 2.2 together establish that $\hat{V}_1^{\ell}(x;X_1)$ is a piece-wise increasing concave function in x. As a result, for each resource $\ell\in\mathcal{L}(X_1)$, we know that $\sum_{x=1}^{x_{\ell,1}}\Delta\hat{V}_1^{l}(x;X_1)\geq x_{\ell,1}\cdot\Delta\hat{V}_1^{\ell}(x_{\ell,1};X_1)$

A.5 Proof of Lemma 2.6

First, by rearranging the expression for the value functions given in (3) for arbitrary time period $t\in[T]$ and inventory vector X_1 we get that

$$\begin{aligned}
\hat{V}_t(X_1;X_1) - \hat{V}_{t+1}(X_1;X_1) &= \sum_{\ell\in\mathcal{L}(X_t)} \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\left[r_{\ell,b} - (\hat{V}_{t+1}(X_1;X_1) - \hat{V}_{t+1}(X_1-e_\ell;X_1))\right]^+ \\
&= \sum_{\ell\in\mathcal{L}(X_t)} \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\left[r_{\ell,b} - \Delta\hat{V}_{t+1}^{\ell}(x_{\ell,1};X_1))\right]^+ \\
&\geq \sum_{\ell\in\mathcal{L}(X_t)} \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\left(r_{\ell,b} - \Delta\hat{V}_{t+1}^{\ell}(x_{\ell,1};X_1))\right).
\end{aligned}$$

Next, summing both sides over all time periods $t\in[T]$ yields

$$\begin{aligned}
\sum_{t\in[T]}\left(\hat{V}_t(X_1;X_1) - \hat{V}_{t+1}(X_1;X_1)\right) &\geq \sum_{t\in[T]}\sum_{\ell\in\mathcal{L}(X_t)} \sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\left(r_{\ell,b} - \Delta\hat{V}_{t+1}^{\ell}(x_{\ell,1};X_1))\right) \\
&= \sum_{t\in[T]}\sum_{b\in\mathcal{B}} p_{b,t}\cdot\max_{\ell\in\mathcal{L}(X_1)} r_{\ell,b} - \sum_{\ell\in\mathcal{L}(X_1)}\left(\sum_{t\in[T]}\sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\cdot\Delta\hat{V}_{t+1}^{l}(x_{\ell,1};X_1)\right) \\
&\geq \sum_{t\in[T]}\sum_{b\in\mathcal{B}} p_{b,t}\cdot\max_{\ell\in\mathcal{L}(X_1)} r_{\ell,b} - \sum_{\ell\in\mathcal{L}(X_1)}\left(\sum_{t\in[T]}\sum_{b\in\mathcal{B}_\ell(X_1)} p_{b,t}\right)\cdot\Delta\hat{V}_1^{\ell}(x_{\ell,1};X_1),
\end{aligned}$$

where the equality follows by definition of $\mathcal{B}_\ell(X_1)$, and the second inequality follows by property (iii) in Claim 2.2. Finally, noting that $\sum_{t\in[T]}\left(\hat{V}_t(X_1) - \hat{V}_{t+1}(X_1)\right) = \hat{V}_1(X_1)$, gives the desired result.

A.6 Proof of Claim 2.7

The problem of computing X_C^* can be shown to be exactly the classical problem of maximizing floats in bank accounts that was originally studied by Cornuejols et al. (1977). In what follows, we formally define the problem of maximizing floats in bank accounts and show that it generalizes the problem of finding X_C^*, which in turn allows us to employ a well-known $(1-\frac{1}{e})$-approximation for the problem of maximizing floats in bank accounts.

Maximizing bank floats. In this problem, we wish to open C bank accounts so as to maximize our float [7]. Let $\mathcal{L}$ be the set of candidate banks that can be opened and let $\mathcal{B}$ be the set of payees who each must be assigned to an opened bank. Furthermore, let $v_{\ell,b}$ be the value of the float created by assigning payee $b \in \mathcal{B}$ to bank $l \in \mathcal{B}$. Clearly, each payee will be matched to the open bank that maximizes $v_{\ell,b}$. Hence we wish to find a subset of bank $S \subseteq \mathcal{L}$ to open such that $|S| \le C$ that maximizes

$$v(S) = \sum_{b\in\mathcal{B}} \max_{l\in S} v_{\ell,b}.$$

To see the connection to the inventory selection problem, note that when $W = \infty$, specifying a subset of resources $\mathcal{L}(X_C^*)$ is equivalent to specifying X_C^*. As a result, the problem of finding X_C^* can be restated as one of choosing a subset of resources $S \subseteq \mathcal{L}$ of cardinality at most C that maximizes the function

$$\sum_{t\in[T]}\sum_{b\in\mathcal{B}} p_{b,t}\cdot \max_{l\in S} r_{\ell,b} = \sum_{b\in\mathcal{B}} \max_{l\in S}\left\{ r_{\ell,b}\cdot \sum_{t\in[T]} p_{b,t}\right\}.$$

Hence by simply setting $v_{\ell,b} = r_{\ell,b}\cdot\sum_{t\in[T]} p_{b,t}$, the reduction becomes clear. Fortunately, Cornuejols et al. (1977) show that the set function $v(S)$ defined above is both monotone and submodular and hence the classical result of G. L. Nemhauser and Fisher (1978) for maximizing monotone submodular set functions immediately yields a $(1-\frac{1}{e})$-approximation. Furthermore, the algorithm to achieve this guarantee is a simple greedy algorithm that adds the bank account (resource) that leads to greatest improvement in float (reward) in each iteration until there is either no improvement or C bank accounts (resources) have been selected.

[7] the technical meaning of float in not essential for defining the problem

A.7 Proof of Lemma 2.8

First, we show that $\bar{X}_1 \in \mathcal{F}(\lfloor \mathcal{P}_{\text{total}} \rfloor + C, C)$. We clearly have that $|\mathcal{L}(\bar{X}_1)| \leq C$, since $\mathcal{L}(\bar{X}_1) = \mathcal{L}(X'_C)$ by construction. Furthermore, the total number of units stocked is

$$\begin{aligned}
\sum_{\ell \in \mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} &= \sum_{\ell \in \mathcal{L}(\bar{X}_1)} \left(\lceil \sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(X'_C)} p_{b,t} \rceil \right) \\
&\leq \sum_{\ell \in \mathcal{L}(\bar{X}_1)} \left(1 + \lfloor \sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(X'_C)} p_{b,t} \rfloor \right) \\
&\leq C + \lfloor \sum_{\ell \in \mathcal{L}(\bar{X}_1)} \sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(X'_C)} p_{b,t} \rfloor \\
&= C + \lfloor \mathcal{P}_{\text{total}} \rfloor
\end{aligned}$$

where the second inequality results since $|\mathcal{L}(\bar{X}_1)| \leq C$.

Next, using Lemma 2.5, we have that

$$\hat{V}_1(\bar{X}_1) \geq \sum_{\ell \in \mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} \cdot \Delta \hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,).$$

Furthermore, using Lemma 2.6, we also have that

$$\begin{aligned}
\hat{V}_1(\bar{X}_1) &\geq \sum_{t \in [T]} \sum_{b \in \mathcal{B}} p_{b,t} \cdot \max_{\ell \in \mathcal{L}(\bar{X}_1)} r_{\ell,b} - \sum_{\ell \in \mathcal{L}(\bar{X}_1)} \left(\sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(\bar{X}_1)} p_{b,t} \right) \cdot \Delta \hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,) \\
&\geq \frac{e-1}{e} \cdot Z^*(\infty, C) - \sum_{\ell \in \mathcal{L}(\bar{X}_1)} \left(\sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(\bar{X}_1)} p_{b,t} \right) \cdot \Delta \hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,),
\end{aligned}$$

where the second inequality follows since $\mathcal{L}(\bar{X}_1) = \mathcal{L}(X'_C)$ and also by using Claim 2.7. Combining the two lower bounds we get that

$$\begin{aligned}
\hat{V}_1(\bar{X}_1) \geq \max \Big\{ &\sum_{\ell \in \mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} \cdot \Delta \hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,), \\
&\frac{e-1}{e} \cdot Z^*(\infty, C) - \sum_{\ell \in \mathcal{L}(\bar{X}_1)} \left(\sum_{t \in [T]} \sum_{b \in \mathcal{B}_\ell(\bar{X}_1)} p_{b,t} \right) \cdot \Delta \hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,) \Big\} \\
\geq \max \Big\{ &\sum_{\ell \in \mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} \cdot \Delta \hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,), \\
&\frac{e-1}{e} \cdot Z^*(\infty, C) - \sum_{\ell \in \mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} \cdot \Delta \hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,) \Big\},
\end{aligned}$$

where the inequality results since $\bar{x}_{\ell,1} = \lceil \sum_{t\in[T]} \sum_{b\in\mathcal{B}_\ell(X'_C)} p_{b,t} \rceil \geq \sum_{t\in[T]} \sum_{b\in\mathcal{B}_\ell(X'_C)} p_{b,t}$. Finally, letting $\alpha \cdot Z^*(\infty, C) = \sum_{\ell\in\mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} \cdot \Delta\hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,)$ for some $\alpha \geq 0$, we get that

$$
\begin{aligned}
\hat{V}_1(\bar{X}_1) &\geq \max\left\{\alpha \cdot Z^*(\infty, C)\ ,\ \frac{e-1}{e} \cdot Z^*(\infty, C) - \alpha \cdot Z^*(\infty, C)\right\} \\
&\geq \min_{\alpha\geq 0} \max\left\{\alpha \cdot Z^*(\infty, C)\ ,\ \frac{e-1}{e} \cdot Z^*(\infty, C) - \alpha \cdot Z^*(\infty, C)\right\} \\
&= \frac{e-1}{2e} \cdot Z^*(\infty, C).
\end{aligned}
$$

A.8 Proof of Lemma 2.9

We have that

$$
\begin{aligned}
\hat{V}_1(X_1^{\ll}) &= \sum_{l\in\mathcal{L}(X_1^{\ll})} \hat{V}_1^\ell(X_1^{\ll}, x_{l,1}^{\ll}) \\
&\geq \sum_{\ell\in\mathcal{L}(\bar{X}_1)} \hat{V}_1^\ell(\bar{X}_1, x_{l,1}^{\ll}) \\
&= \sum_{\ell\in\mathcal{L}(\bar{X}_1)} \hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,) - \sum_{\ell\in\mathcal{L}(\bar{X}_1)} \left(\hat{V}_1^\ell(\bar{x}_{\ell,1}; \bar{X}_1,) - \hat{V}_1^\ell(\bar{X}_1, x_{l,1}^{\ll}).\right) \\
&= \sum_{\ell\in\mathcal{L}(\bar{X}_1)} \sum_{x=0}^{\bar{x}_{\ell,1}} \Delta\hat{V}_1^\ell(\bar{X}_1, x) - \sum_{\ell\in\mathcal{L}(\bar{X}_1)} \sum_{x=x_{l,1}^{\ll}+1}^{\bar{x}_{\ell,1}} \Delta\hat{V}_1^\ell(\bar{X}_1, x) \\
&\geq \hat{V}_1(\bar{X}_1) - \frac{[\sum_{\ell\in\mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} - W]^+}{\sum_{\ell\in\mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1}} \cdot \hat{V}_1(\bar{X}_1) \\
&\geq \min\{1, \frac{W}{\sum_{\ell\in\mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1}}\} \cdot \hat{V}_1(\bar{X}_1) \\
&\geq \min\{1, \frac{W}{\lfloor\mathcal{P}_{\text{total}}\rfloor + C}\} \cdot \hat{V}_1(\bar{X}_1) \\
&\geq \frac{e-1}{2e} \cdot \min\{1, \frac{W}{\lfloor\mathcal{P}_{\text{total}}\rfloor + C}\} \cdot Z^*(\infty, C).
\end{aligned}
$$

The first inequality holds immediately by first noting that if $l \in \mathcal{L}(\bar{X}_1) \setminus \mathcal{L}(X_1^{\ll})$, then $x_{l,1}^{\ll} = 0$ and then by applying Claim 2.3, since $\mathcal{L}(X_1^{\ll}) \subseteq \mathcal{L}(\bar{X}_1)$. The second inequality uses the fact that the marginal value of each resource is decreasing in the inventory level under our value functions (property (ii) of Claim 2.2) and the fact that Algorithm 1 removes the $[\sum_{\ell\in\mathcal{L}(\bar{X}_1)} \bar{x}_{\ell,1} - W]^+$ least profitable units of inventory from $\bar{X}_1$. The last inequalities follow due to conditions (i) and (ii) of Lemma 2.8 respectively.

A.9 Proof of Lemma 2.10

We first construct an intermediate solution $(\hat{y}, \hat{x})$, which stocks at most $W + C$ units of inventory, as follows. We set

$$\hat{y}^t_{\ell,b} = \begin{cases} \sum_{\ell' \in \mathcal{L}(x^*)} y^{*t}_{\ell',b}, & \text{if } b \in \mathcal{B}_l(x^*) \\ 0, & \text{otherwise,} \end{cases}$$

essentially shifting all matched demand under y^* to its ideal resource among those stocked under x^*. Next, we set

$$\hat{x}_l = \lceil \sum_{t \in [T]} \sum_{b \in \mathcal{B}} \hat{y}^t_{\ell,b} \rceil,$$

and note that

$$\begin{aligned} \sum_{\ell \in \mathcal{L}} \hat{x}_\ell &\leq \sum_{t \in [T]} \sum_{\ell \in \mathcal{L}} \sum_{b \in \mathcal{B}} \hat{y}^t_{\ell,b} + C \\ &= \sum_{t \in [T]} \sum_{\ell \in \mathcal{L}} \sum_{b \in \mathcal{B}} y^{*t}_{\ell,b} + C \\ &\leq W + C. \end{aligned}$$

The first inequality above follows from the definition of $\hat{x}_l$ and the fact that we stock at most C resources. From here, we sequentially remove the least valuable units of inventory, and unmatch the demand assigned to these units, until we arrive at starting inventory vector that has at most W units. The exact nature of this greedy removal step is given in Algorithm 2, where Worst_Unit$_\ell$ is the reward of the least valuable unit of ℓ that is in-use under the solution $\bar{y}$. Within the definition of Worst_Unit$_\ell$, the constraint $\sum_{t \in [T]} \sum_{b \in \mathcal{B}} \bar{y}^t_{\ell,b} z^t_{\ell,b} = 1 - (\bar{x}_\ell - \sum_{t \in [T]} \sum_{b \in \mathcal{B}} \bar{y}^t_{\ell,b})$ ensures that after each iteration of the while loop, we are left with an integer number of units of resource ℓ^*.

After carrying out Algorithm 2 to yield $\bar{x}, \bar{y}$, we simply set $\bar{z} = z^*$. It is not difficult to see that the solution $(\bar{y}, \bar{x}, \bar{z})$ is feasible to Fluid-IP based on its construction. Moreover, since $\hat{y}^t_{\ell,b} > 0$ only if $b \in \mathcal{B}_l(x^*)$ and $\bar{x}_\ell > 0$ by construction, we know that condition (i) of the lemma statement is satisfied. To see condition (ii), note that

$$\begin{aligned} \sum_{t \in [T]} \sum_{\ell \in \mathcal{L}} \sum_{b \in \mathcal{B}} r_{\ell,b} \bar{y}^t_{\ell,b} &\geq \frac{W}{W + C} \cdot \sum_{t \in [T]} \sum_{\ell \in \mathcal{L}} \sum_{b \in \mathcal{B}} r_{\ell,b} \hat{y}^t_{\ell,b} \\ &\geq \frac{1}{2} \cdot \sum_{t \in [T]} \sum_{\ell \in \mathcal{L}} \sum_{b \in \mathcal{B}} r_{\ell,b} \hat{y}^t_{\ell,b} \\ &\geq \frac{1}{2} \cdot \text{Fluid}(W, C). \end{aligned}$$

The first inequality follows because we remove at most the C least valuable resources, and the

Algorithm 2 Building $\bar{x}, \bar{y}$

procedure GreedyRemoveDLP($\hat{x}, \hat{y}$)
  $\bar{x} \leftarrow \hat{x}$
  $\bar{y} \leftarrow \hat{y}$
  while $\sum_{\ell \in \mathcal{L}} \bar{x}_\ell > W$ **do**
    for $\ell \in \mathcal{L}(\bar{x})$ **do**
      $\text{Worst\_Unit}_\ell = \min_{z^t_{\ell,b} \in [0,1]} \{ \sum_{t \in [T]} \sum_{b \in \mathcal{B}} r_{\ell,b} \bar{y}^t_{\ell,b} z^t_{\ell,b} : \sum_{t \in [T]} \sum_{b \in \mathcal{B}} \bar{y}^t_{\ell,b} z^t_{\ell,b} = 1 - (\bar{x}_\ell - \sum_{t \in [T]} \sum_{b \in \mathcal{B}} \bar{y}^t_{\ell,b}) \}$
    end for
    $\ell^* \leftarrow \operatorname{argmin}_{\ell \in \mathcal{L}(\bar{x})} \text{Worst\_Unit}_\ell$ (Breaking ties arbitrarily)
    $\bar{y}^t_{\ell^*,b} = \bar{y}^t_{\ell^*,b} \cdot (1 - z^t_{\ell^*,b})$
    $\bar{x}_{\ell^*} = \bar{x}_{\ell^*} - 1$
  end while
  return $\bar{x}, \bar{y}$
end procedure

second holds because $W \geq C$. The final inequality holds since $\hat{y}$ shifts all matched demand under y^* to its ideal resource, thus it must achieve an objective of at least $\text{Fluid}(W, C)$.

A.10 Proof of Lemma 2.11

We prove this result by establishing the following two lower bounds on $\hat{V}_1(\bar{x})$.

Lower Bound 1. Exactly mirroring the bound established in Lemma 2.5, we have

$$\hat{V}_1(\bar{x}) \geq \sum_{l \in \mathcal{L}(\bar{x})} \bar{x}_\ell \cdot \Delta \hat{V}^\ell_1(\bar{x}_\ell; \bar{x}).$$

Lower Bound 2. Closely resembling the bound established in Lemma 2.6, we have

$$\hat{V}_1(\bar{x}) \geq \sum_{t \in [T]} \sum_{\ell \in \mathcal{L}} \sum_{b \in \mathcal{B}} r_{\ell,b} \bar{y}^t_{\ell,b} - \sum_{l \in \mathcal{L}(\bar{x})} \bar{x}_\ell \cdot \Delta \hat{V}^\ell_1(\bar{x}_\ell; \bar{x}).$$

We prove this second bound at the end of this section, but first note that combining both bounds yields

$$\begin{aligned}
\hat{V}_1(\bar{x}) &\geq \max \{\text{Lower Bound 1}, \text{Lower Bound 2}\} \\
&\geq \frac{1}{2} \cdot \sum_{t \in [T]} \sum_{\ell \in \mathcal{L}} \sum_{b \in \mathcal{B}} r_{\ell,b} \bar{y}^t_{\ell,b} \\
&\geq \frac{1}{4} \cdot \text{Fluid}(W, C) \\
&\geq \frac{1}{4} \cdot Z^*(W, C),
\end{aligned}$$

where the third inequality follows from condition (ii) of Lemma 2.10.

Proof of Lower Bound 2. Following the proof of Lemma 2.6, we see that

$$\begin{aligned}
\hat{V}_t(\bar{x};\bar{x}) - \hat{V}_{t+1}(\bar{x};\bar{x}) &= \sum_{\ell\in\mathcal{L}(\bar{x})}\sum_{b\in\mathcal{B}_\ell(\bar{x})} p_{b,t}\cdot\left[r_{\ell,b} - \Delta\hat{V}^{\ell}_{t+1}(\bar{x}_\ell;\bar{x}))\right]^+ \\
&\geq \sum_{l\in\mathcal{L}(\bar{x})}\sum_{b\in\mathcal{B}_\ell(\bar{x})} \bar{y}^t_{\ell,b}\cdot\left[r_{\ell,b} - \Delta\hat{V}^{\ell}_{t+1}(\bar{x}_\ell;\bar{x}))\right]^+ \\
&\geq \sum_{l\in\mathcal{L}(\bar{x})}\sum_{b\in\mathcal{B}_\ell(\bar{x})} \bar{y}^t_{\ell,b}\cdot\left(r_{\ell,b} - \Delta\hat{V}^{\ell}_{t+1}(\bar{x}_\ell;\bar{x}))\right),
\end{aligned}$$

where the first inequality follows because $\bar{y}^t_{\ell,b} \leq p_{b,t}$ as stated in the first constraint of Fluid-IP. Next, summing both sides over all time periods $t \in [T]$ yields

$$\begin{aligned}
\sum_{t\in[T]}\left(\hat{V}_t(\bar{x};\bar{x}) - \hat{V}_{t+1}(\bar{x};\bar{x}\right) &\geq \sum_{t\in[T]}\sum_{l\in\mathcal{L}(\bar{x})}\sum_{b\in\mathcal{B}_\ell(\bar{x})} \bar{y}^t_{\ell,b}\cdot\left(r_{\ell,b} - \Delta\hat{V}^{\ell}_{t+1}(\bar{x}_\ell;\bar{x}))\right) \\
&= \sum_{t\in[T]}\sum_{l\in\mathcal{L}(\bar{x})}\sum_{b\in\mathcal{B}_\ell(\bar{x})} \bar{y}^t_{\ell,b}\cdot r_{\ell,b} - \sum_{l\in\mathcal{L}(\bar{x})}\left(\sum_{t\in[T]}\sum_{b\in\mathcal{B}_\ell(\bar{x})} \bar{y}^t_{\ell,b}\cdot\Delta\hat{V}^{l}_{t+1}(\bar{x}_\ell;\bar{x})\right) \\
&\geq \sum_{t\in[T]}\sum_{l\in\mathcal{L}(\bar{x})}\sum_{b\in\mathcal{B}_\ell(\bar{x})} \bar{y}^t_{\ell,b}\cdot r_{\ell,b} - \sum_{l\in\mathcal{L}(\bar{x})}\left(\sum_{t\in[T]}\sum_{b\in\mathcal{B}_\ell(\bar{x})} \bar{y}^t_{\ell,b}\right)\cdot\Delta\hat{V}^{l}_{1}(\bar{x}_\ell;\bar{x}) \\
&\geq \sum_{t\in[T]}\sum_{l\in\mathcal{L}(\bar{x})}\sum_{b\in\mathcal{B}_\ell(\bar{x})} \bar{y}^t_{\ell,b}\cdot r_{\ell,b} - \sum_{l\in\mathcal{L}(\bar{x})} \bar{x}_\ell\cdot\Delta\hat{V}^{l}_{1}(\bar{x}_\ell;\bar{x}),
\end{aligned}$$

where the last inequality follows by the second constraint of Fluid-IP. Finally, noting that $\sum_{t\in[T]}\left(\hat{V}_t(\bar{x};\bar{x}) - \hat{V}_{t+1}(\bar{x};\bar{x}\right) = \hat{V}_1(\bar{x})$ yields the desired result.

B Proofs from Section 4

B.1 Proof of Lemma 4.1

Intermediate results. Before proceeding with the proof of Lemma 4.1, we first present two intermediate claims, whose proofs can each be found at the end of this section. In particular, the first claim bounds the differences in matching rewards between any two trailers, and is used to prove the second claim, which bounds differences of the optimal value functions for inventory states that differ by a single unit.

Claim B.1. *For arbitrary truck $b \in \mathcal{B}$, and trailers $\ell, \ell^+ \in \mathcal{L}$ that satisfy $w_{\ell^+} > w_\ell$, we have that*

$$c\cdot(w_\ell - w_{\ell^+}) \leq r_{\ell^+,b} - r_{\ell,b} \leq r\cdot(w_{\ell^+} - w_\ell).$$

Claim B.2. *For any time period $t \in [T]$, inventory vector X_t, and trailers $\ell, \ell^+ \in \mathcal{L}(X_t)$ that satisfy $w_{\ell^+} > w_\ell$, we have that* *(i)* $V_t(X_t - e_\ell) - V_t(X_t - e_{\ell^+}) \geq c \cdot (w_\ell - w_{\ell^+})$

(ii) $V_t(X_t - e_\ell) - V_t(X_t - e_{\ell^+}) \leq r \cdot (w_{\ell^+} - w_\ell)$.

Proof of Lemma 4.1. To begin, let

$$\ell^* = \underset{\ell \in \mathcal{L}(X_t)}{\operatorname{argmax}} \left\{ r_{\ell,b} + V_{t+1}(X_t - e_\ell) \right\},$$

and assume by way of contradiction that

$$r_{\ell^*,b} + V_{t+1}(X_t - e_{\ell^*}) > \max_{\ell \in \{\ell^\uparrow(b,X_t), \ell^\downarrow(b,X_t)\}} \left\{ r_{\ell,b} + V_{t+1}(X_t - e_\ell) \right\}. \tag{11}$$

We show that (11) cannot hold by considering the following two cases, where for ease of notation, we let $\ell^\uparrow = \ell^\uparrow(b, X_t)$ and $\ell^\downarrow = \ell^\downarrow(b, X_t)$.

- Case 1 - $w_{\ell^*} > w_{\ell^\uparrow}$: In this case, we will show that

  $$r_{\ell^*,b} + V_{t+1}(X_t - e_{\ell^*}) \leq r_{\ell^\uparrow,b} + V_{t+1}(X_t - e_{\ell^\uparrow}),$$

  which contradicts (11). To do so, note that

  $$\begin{aligned}
  &(r_{\ell^*,b} - r_{\ell^\uparrow,b}) + (V_{t+1}(X_t - e_{\ell^*}) - V_{t+1}(X_t - e_{\ell^\uparrow})) \\
  &\quad = c \cdot (w_{\ell^\uparrow} - w_{\ell^*}) + (V_{t+1}(X_t - e_{\ell^*}) - V_{t+1}(X_t - e_{\ell^\uparrow})) \\
  &\quad \leq c \cdot (w_{\ell^\uparrow} - w_{\ell^*}) - c \cdot (w_{\ell^*} - w_{\ell^\uparrow}) = 0.
  \end{aligned}$$

  The first equality uses the structure of the matching rewards given in (7), along with the definition of the trailer $\ell^\uparrow$. The second inequality applies property (ii) of Claim B.2, noting that $w_{\ell^*} > w_{\ell^\uparrow}$.

- Case 2 - $w_{\ell^*} < w_{\ell^\downarrow}$: In this case, we will show that

  $$r_{\ell^*,b} + V_{t+1}(X_t - e_{\ell^*}) \leq r_{\ell^\downarrow,b} + V_{t+1}(X_t - e_{\ell^\downarrow}),$$

  which contradicts (11). To do so, note that

  $$\begin{aligned}
  &(r_{\ell^*,b} - r_{\ell^\downarrow,b}) + (V_{t+1}(X_t - e_{\ell^*}) - V_{t+1}(X_t - e_{\ell^\downarrow})) \\
  &\quad = -r \cdot (w_{\ell^\downarrow} - w_{\ell^*}) + (V_{t+1}(X_t - e_{\ell^*}) - V_{t+1}(X_t - e_{\ell^\uparrow})) \\
  &\quad \leq -r \cdot (w_{\ell^\downarrow} + w_{\ell^*}) + r \cdot (w_{\ell^\downarrow} + w_{\ell^*}) = 0.
  \end{aligned}$$

  The first equality uses the structure of the matching rewards given in (7), along with the definition of the trailer $\ell^\uparrow$. The second inequality applies property (ii) of Claim B.2, noting that $w_{\ell^*} < w_{\ell^\downarrow}$.

Proof of Claim B.1. To establish the lower and upper bounds, we consider the following three cases that are based on whether the assignment of trailer b to either ℓ^+ or ℓ incurs scaleback costs.

- Case 1 - Neither trailer incurs scaleback costs: In this case, we have that $r_{\ell^+,b} - r_{\ell,b} = r \cdot (w_{\ell^+} - w_\ell) > 0$, where the inequality holds since $w_{\ell^+} > w_\ell$. Noting that $c \cdot (w_\ell - w_{\ell^+}) < 0$, we see that the lower bound holds as well.

- Case 2 - Both trailer incurs scaleback costs: In this case, we have that $r_{\ell^+,b} - r_{\ell,b} = c \cdot (w_\ell - w_{\ell^+}) < 0$. Noting that $r \cdot (w_{\ell^+} - w_\ell) > 0$, we see that the upper bound holds as well.

- Case 3 - Only trailer ℓ^+ incurs scaleback costs: In this case, we establish the upper bound by noting that

$$\begin{aligned} r_{\ell^+,b} - r_{\ell,b} &= ((80,000 - \Omega_b) \cdot r - ((w_{\ell^+} + \Omega_b) - 80,000) \cdot c) - rw_\ell \\ &\leq (80,000 - \Omega_b) \cdot r - rw_\ell \\ &\leq r \cdot (w_{\ell^+} - w_\ell). \end{aligned}$$

  where the last inequality follows because $w_{\ell^+} + \Omega_b \geq 80,000$, since trailer ℓ^+ incurs a scaleback costs. To establish the lower bound, we proceed as follows

$$\begin{aligned} r_{\ell^+,b} - r_{\ell,b} &= ((80,000 - \Omega_b) \cdot r - ((w_{\ell^+} + \Omega_b) - 80,000) \cdot c) - rw_\ell \\ &\geq (80,000 - (\Omega_b + w_{\ell^+})) \cdot c \\ &\geq c \cdot (w_\ell - w_{\ell^+}). \end{aligned}$$

  The first inequality follows because $w_\ell + \Omega_b \leq 80,000$, since trailer ℓ does not incur scaleback costs, and so $(80,000 - \Omega_b) \cdot r - rw_\ell \geq 0$. The second inequality directly uses the fact that $w_\ell + \Omega_b \leq 80,000$.

Proof of property (i) of Claim B.2. We will prove this lower bound via induction over t. Since $V_{T+1}(\cdot) = 0$, the base case of $t = T + 1$ holds trivially, and so we move to establishing the general case of $t \in [T]$. Recalling that

$$V_t(X_t) = \sum_{b \in \mathcal{B}} p_{b,t} \cdot \max_{\ell \in \mathcal{L}(X_t)} \{r_{\ell,b} + V_{t+1}(X_t - e_\ell)\} + \left(1 - \sum_{b \in \mathcal{B}} p_{b,t}\right) \cdot V_{t+1}(X_t),$$

we let

$$\begin{aligned} k_b &= \operatorname*{argmax}_{k \in \mathcal{L}(X_t - e_\ell)} \{r_{k,b} + V_{t+1}(X_t - e_\ell - e_k)\} \\ k_b^+ &= \operatorname*{argmax}_{k \in \mathcal{L}(X_t - e_{\ell^+})} \{r_{k,b} + V_{t+1}(X_t - e_{\ell^+} - e_k)\} \end{aligned}$$

denote the optimal trailers assigned under period-t inventory levels of $X_t - e_\ell$ and $X_t - e_{\ell^+}$ respectively, given the arrival of a truck of type b. Furthermore, we define

$$\Delta_b = (r_{k_b,b} + V_{t+1}(X_t - e_\ell - e_{k_b})) - (r_{k_b^+,b} + V_{t+1}(X_t - e_{\ell^+} - e_{k_b^+})).$$

We will prove that, for any $b \in \mathcal{B}$, we have that $\Delta_b \geq c \cdot (w_\ell - w_{\ell^+})$, in which case we get that

$$\begin{aligned} V_t(X_t - e_\ell) - V_t(X_t - e_{\ell^+}) &= \sum_{b\in\mathcal{B}} p_{b,t} \cdot \Delta_b + \left(1 - \sum_{b\in\mathcal{B}} p_{b,t}\right) \cdot (V_{t+1}(X_t - e_\ell) - V_{t+1}(X_t - e_{\ell^+})) \\ &\geq \sum_{b\in\mathcal{B}} p_{b,t} \cdot \Delta_b + \left(1 - \sum_{b\in\mathcal{B}} p_{b,t}\right) \cdot c \cdot (w_\ell - w_{\ell^+}) \\ &\geq c \cdot (w_\ell - w_{\ell^+}), \end{aligned}$$

where the first inequality follows by the induction hypothesis.

To show that, for any $b \in \mathcal{B}$, we have that $\Delta_b \geq c \cdot (w_\ell - w_{\ell^+})$, we consider the following two cases.

- Case 1 - $k_b^+ \in \mathcal{L}(X_t - e_\ell)$: In this case, we observe that

$$r_{k_b,b} + V_{t+1}(X_t - e_\ell - e_{k_b}) \geq r_{k_b^+,b} + V_{t+1}(X_t - e_\ell - e_{k_b^+}) \tag{12}$$

  by the optimality of the trailer k_b. From here, observe that

$$\begin{aligned} \Delta_b &= (r_{k_b,b} + V_{t+1}(X_t - e_\ell - e_{k_b})) - (r_{k_b^+,b} + V_{t+1}(X_t - e_{\ell^+} - e_{k_b^+})) \\ &\qquad + V_{t+1}(X_t - e_\ell - e_{k_b^+}) - V_{t+1}(X_t - e_\ell - e_{k_b^+}) \\ &= (r_{k_b,b} + V_{t+1}(X_t - e_\ell - e_{k_b})) - (r_{k_b^+,b} + V_{t+1}(X_t - e_\ell - e_{k_b^+})) \\ &\qquad + V_{t+1}(X_t - e_\ell - e_{k_b^+}) - V_{t+1}(X_t - e_{\ell^+} - e_{k_b^+}) \\ &\geq V_{t+1}(X_t - e_\ell - e_{k_b^+}) - V_{t+1}(X_t - e_{\ell^+} - e_{k_b^+}) \\ &\geq c \cdot (w_\ell - w_{\ell^+}), \end{aligned}$$

  where the first inequality follows by (12) and the second inequality follows by the induction hypothesis.

- Case 2 - $k_b^+ \notin \mathcal{L}(X_t - e_\ell)$: In this case, we must have that $k_b^+ = \ell$, and so

$$\begin{aligned} \Delta_b &= (r_{k_b,b} + V_{t+1}(X_t - e_\ell - e_{k_b})) - (r_{\ell,b} + V_{t+1}(X_t - e_{\ell^+} - e_\ell)) \\ &\geq (r_{\ell^+,b} + V_{t+1}(X_t - e_\ell - e_{\ell^+})) - (r_{\ell,b} + V_{t+1}(X_t - e_{\ell^+} - e_\ell)) \\ &= r_{\ell^+,b} - r_{\ell,b} + (V_{t+1}(X_t - e_\ell - e_{\ell^+}) - V_{t+1}(X_t - e_{\ell^+} - e_\ell)) \\ &\geq c \cdot (w_\ell - w_{\ell^+}), \end{aligned}$$

where the first inequality follows by the optimality of the trailer of k_b and the fact that $\ell^+ \in \mathcal{L}(X_t - e_\ell)$, and the second inequality uses Claim B.1.

Proof of property (ii) of Claim B.2. We will also prove the upper bound via induction over t. Since $V_{T+1}(\cdot) = 0$, the base case of $t = T+1$ holds trivially, and so we move to establishing the general case of $t \in [T]$. Borrowing notation from the proof of part (i), we will show that, for any $b \in \mathcal{B}$, we have that $\Delta_b \le r \cdot (w_{\ell^+} - w_\ell)$, in which case we get that

$$\begin{aligned} V_t(X_t - e_\ell) - V_t(X_t - e_{\ell^+}) &= \sum_{b \in \mathcal{B}} p_{b,t} \cdot \Delta_b + \left(1 - \sum_{b \in \mathcal{B}} p_{b,t}\right) \cdot (V_{t+1}(X_t - e_\ell) - V_{t+1}(X_t - e_{\ell^+})) \\ &\le \sum_{b \in \mathcal{B}} p_{b,t} \cdot \Delta_b + \left(1 - \sum_{b \in \mathcal{B}} p_{b,t}\right) \cdot r \cdot (w_{\ell^+} - w_\ell) \\ &\le r \cdot (w_{\ell^+} - w_\ell), \end{aligned}$$

where the first inequality follows by the induction hypothesis.

To show that, for any $b \in \mathcal{B}$, we have that $\Delta_b \le r \cdot (w_{\ell^+} - w_\ell)$, we again consider the following two cases.

- Case 1 - $k_b \in \mathcal{L}(X_t - e_{\ell^+})$: In this case, we observe that

$$r_{k_b,b} + V_{t+1}(X_t - e_{\ell^+} - e_{k_b}) \le r_{k_b^+,b} + V_{t+1}(X_t - e_{\ell^+} - e_{k_b^+}) \tag{13}$$

  by the optimality of the trailer k_b^+. From here, observe that

$$\begin{aligned} \Delta_b &= (r_{k_b,b} + V_{t+1}(X_t - e_\ell - e_{k_b})) - (r_{k_b^+,b} + V_{t+1}(X_t - e_{\ell^+} - e_{k_b^+})) \\ &\qquad + V_{t+1}(X_t - e_{\ell^+} - e_{k_b}) - V_{t+1}(X_t - e_{\ell^+} - e_{k_b}) \\ &= (r_{k_b,b} + V_{t+1}(X_t - e_{\ell^+} - e_{k_b})) - (r_{k_b^+,b} + V_{t+1}(X_t - e_{\ell^+} - e_{k_b^+})) \\ &\qquad + V_{t+1}(X_t - e_\ell - e_{k_b}) - V_{t+1}(X_t - e_{\ell^+} - e_{k_b}) \\ &\le V_{t+1}(X_t - e_\ell - e_{k_b}) - V_{t+1}(X_t - e_{\ell^+} - e_{k_b}) \\ &\le r \cdot (w_{\ell^+} - w_\ell), \end{aligned}$$

  where the first inequality follows by (13) and the second inequality follows by the induction hypothesis.

- Case 2 - $k_b \notin \mathcal{L}(X_t - e_{\ell^+})$: In this case, we must have that $k_b = \ell^+$, and so

$$\begin{aligned}\Delta_b &= (r_{\ell^+,b} + V_{t+1}(X_t - e_\ell - e_{\ell^+})) - (r_{k_b^+,b} + V_{t+1}(X_t - e_{\ell^+} - e_{k_b^+}))\\ &\leq (r_{\ell^+,b} + V_{t+1}(X_t - e_\ell - e_{\ell^+})) - (r_{\ell,b} + V_{t+1}(X_t - e_{\ell^+} - e_\ell))\\ &= r_{\ell^+,b} - r_{\ell,b} + (V_{t+1}(X_t - e_\ell - e_{\ell^+}) - V_{t+1}(X_t - e_{\ell^+} - e_\ell))\\ &\leq r \cdot (w_{\ell^+} - w_\ell),\end{aligned}$$

where the first inequality follows by the optimality of the trailer of k_b^+ and the fact that $\ell \in \mathcal{L}(X_t - e_{\ell^+})$, and the second inequality uses Claim B.1.

B.2 Proof of Claim 4.2

Under any initial inventory vector $X_1 \in \mathcal{F}(\infty, C)$, it is easy to see that both the optimal policy and the BoR policy reduce to a myopic greedy policy in which each truck is assigned its ideal trailer type, and so

$$V_1(X_1) = \hat{V}_1(X_1) = \sum_{\ell \in \mathcal{L}(X_1)} \sum_{b \in \mathcal{B}_\ell(X_1)} r_{\ell,b} \cdot \left(\sum_{t \in [T]} p_{b,t} \right),$$

Consequently, to prove the claim, it suffices to show that

$$X_C^* = \max_{X_1 \mathcal{F}(\infty,C)} \sum_{\ell \in \mathcal{L}(X_1)} \sum_{b \in \mathcal{B}_\ell(X_1)} r_{\ell,b} \cdot \left(\sum_{t \in [T]} p_{b,t} \right) \tag{14}$$

can be computed in polynomial time. Since $X_C^* \in \mathcal{F}(\infty, C)$, we note that it suffices to specify $\mathcal{L}(X_C^*)$ alone, since each utilized trailer will be prepared with infinite capacity.

In what follows, we present a simple dynamic program that can be used to compute $\mathcal{L}(X_C^*)$ in a running time of $O(L^4 B)$. For the remainder of this section, we assume the trailer types are indexed in increasing order of weight. Furthermore, it will be useful to introduce a dummy trailer type with index 0 that has a weight of $-\infty$. We first observe that, for any $X_1 \in \mathcal{F}(\infty, C)$ and $\ell \in \mathcal{L}(X_1)$, we can fully compute $\mathcal{B}_\ell(X_1)$ with knowledge of only $\ell^+ = \min\{\ell' \in \mathcal{L}(X_1) : w_{\ell'} > w_\ell\}$ and $\ell^- = \max\{\ell' \in \mathcal{L}(X_1) : w_{\ell'} < w_\ell\}$, rather than the entire inventory vector X_1. More formally, let

$$\mathcal{B}_\ell(\ell^-, \ell^+) = \{b \in \mathcal{B} : \ell = \operatorname*{argmax}_{k \in \{\ell, \ell^-, \ell^+\}} r_{k,b}\},$$

and note that the piecewise linear structure of the matching rewards immediately gives that $\mathcal{B}_\ell(X_1) = \mathcal{B}_\ell(\ell^-, \ell^+)$.

With this insight, we propose the following dynamic program to solve (14). The value functions $\mathcal{J}(\ell^-, \ell, c)$ represent the maximum expected reward that can be garnered from trailers $\ell, \dots, L$, given that

- trailer ℓ^- is the highest indexed trailer type used among trailers $1, \dots, \ell - 1$.
- trailer ℓ will be used.
- at most c more trailer type can be selected.

Formally, we have

$$\mathcal{J}(\ell^-, \ell, c) = \max_{\ell^+ \in \{\ell+1, \dots, L+1\}} \sum_{b \in \mathcal{B}_\ell(\ell^-, \ell^+)} r_{\ell,b} \cdot \left(\sum_{t \in [T]} p_{b,t} \right) + \mathcal{J}(\ell, \ell^+, c-1) \tag{15}$$

with base cases $\mathcal{J}(\cdot, L+1, \cdot) = 0$ and $\mathcal{J}(\cdot, \cdot, 0) = 0$. In this case, it is clear that $Z^*(\infty, C) = \max_{\ell \in [L+1]} \mathcal{J}(0, \ell, C-1)$. Moreover, letting $\ell_1^* = \mathrm{argmax}_{\ell \in [L+1]} \mathcal{J}(0, \ell, C-1)$, we will recover $\mathcal{L}(X_C^*)$ by traversing the above dynamic program starting from state $(0, \ell_1^*, C-1)$, while also including trailer type ℓ_1^*.

The dynamic program outlined in (15) has $O(L^2 C) = O(L^3)$ states. Furthermore, the maximization problem can be solved by enumerating over all trailer $\ell^+ \in \{\ell+1, \dots, L+1\}$, of which there are most $O(L)$. Hence the final running time required to recover $\mathcal{L}(X_C^*)$ is $O(L^4 B)$.

C Inapproximability Results

C.1 Inaproximability - reduction to max coverage

In this section, we establish the following inapproximability result.

Theorem C.1. *It is NP-Hard to find an $X_1 \in \mathcal{F}(\infty, C)$ that satisfies*

$$V_1(X_1) \geq \left(\frac{e-1}{e} + \epsilon \right) \cdot Z^*(\infty, C),$$

for every $C \in \mathbb{Z}_+$ and $\epsilon > 0$, unless P=NP.

Theorem C.1 is a direct consequence of Feige (1998), who shows that the maximum coverage problem cannot be approximated within a factor better than $\frac{e-1}{e} + \epsilon$ for any $\epsilon > 0$, unless $P = NP$. In what follows, we show that the problem maximum coverage problem is a special case of our problem, in which the matching rewards are binary and $W = \infty$. This inapproximability result implies that, even when $W = \infty$, no approach can garner an expected reward that exceeds $(\frac{e-1}{e}) \cdot Z^*(\infty, C)$ in general.

The maximum coverage problem. In the maximum coverage problem, we are given a base set of elements $E = \{e_1, e_2, \dots, e_n\}$ and a collection of subsets of these elements $S = \{S_1, \dots, S_m\}$. The goal is to choose $S' \subset S$ that satisfies $|S'| \leq C$ and that maximizes $|\cup_{S_i \in S'} S_i|$, i.e. the number of elements "covered" by S'.

The reduction. The reduction works as follows. We have n resources, one for each subset $S_\ell \in S$, and m customer types, one for each element $e_b \in E$. There are m time periods, one for each element, where during period $b \in [m]$, we have that $p_{b,b} = 1$. The reward for matching customer b to resource ℓ is 1 if $e_b \in S_\ell$, and 0 otherwise. In this case, it is easy to show that

$$\max_{X_1 \in \mathcal{F}(\infty, C)} V_1(X_1) = \max_{S \subseteq [n]: |S| \leq C} \left\{ \sum_{b \in [m]} \min \left\{ 1, \sum_{\ell \in S} \mathbb{1}_{e_b \in S_\ell} \right\} \right\}, \tag{16}$$

where the latter problem is precisely the maximum coverage problem. The important observation needed to obtain (16) is that, when $W = \infty$, we get that $V_1(X_1) = \sum_{t \in [T]} \sum_{b \in \mathcal{B}} p_{b,t} \cdot \max_{\ell \in \mathcal{L}(X_1)} r_{\ell,b}$, since it is trivially optimal to assign each customer type to their ideal resource.

C.2 Inaproximability - no rejections

In this section, we establish the following inapproximability result.

Theorem C.2. *If each customer must be assigned an available resource, then for any $\alpha > 0$, it is NP-Hard to find an $X_1 \in \mathcal{F}(\infty, C)$ that satisfies*

$$\hat{V}_1(X_1) \geq \alpha \cdot Z^*(\infty, C),$$

for every $C \in [T]$, unless P=NP.

We prove the results via a reduction from the vertex cover problem, which is one of Karp's 21 NP-Complete problems (Karp, 1972). The decision version of the vertex cover problem asked whether, for input graph $G = (V, E)$, there exists a subset of vertices $S \subseteq V$ satisfying $|S| \leq k$ such that each edge has an endpoint in S. We assume that the input graph has n vertices indexed $V = \{v_1, \dots, v_n\}$ and m edges index $E = \{e_1, \dots, e_m\}$.

Given an instance of the vertex cover problem, we create an instance of our joint inventory selection and online matching problem as follows:

- We have n resources (one for each vertex) and m customer types (one for each edge).

- We set $C = k$ and $W = \infty$. Since $W = \infty$, the optimal matching policy is to simply to match each arriving customer to her ideal resource, among those that are stocked.

- There are m time periods (one for each edge), where in period $t \in [m]$, customer type t arrives with certainty i.e. $p_{t,t} = 1$.

- The reward for matching type $b \in [m]$ to resource $l \in [n]$ is $r_{\ell,b} = 1$ if v_l is an endpoint of edge e_b, otherwise, we have that $r_{\ell,b} = -2m$.

In the remainder of the proof, we show that for any $X_1 \in \mathcal{F}(\infty, k)$ that satisfies $V_1(X_1) > 0$, the subset of vertices $S = \{v_l \in V : \ell \in \mathcal{L}(X_1)\}$ must be a vertex cover of size k. To show this result, we begin by noting that if $V_1(X_1) > 0$, then we must have that $V_1(X_1) = m$, since if a reward of 1 is not earned in a particular time period, then the only alternative is that we earned a reward of $-2m$ (since we cannot reject arriving customers), which would ensure that $V_1(x_1) < 0$. Finally, noting that $V_1(X_1) = m$ if and only in each period $t \in [m]$, there exists a resource $\ell \in \mathcal{L}(X_1)$ such that v_l is an endpoint of e_t, immediately implies that $S = \{v_l \in V : \ell \in \mathcal{L}(X_1)\}$ is a vertex cover of size at most k.

Summary. First, note that if there exists an inventory vector $X_1 \in \mathcal{F}(\infty, k)$ such that $V_1(X_1) > 0$, then an α-approximation (for any $\alpha > 0$) to our joint inventory selection and online matching problem must yield an inventory vector $\hat{X}_1 \in \mathcal{F}(\infty, k)$ that satisfies $V_1(\hat{X}_1) > 0$. As such, if $V_1(\hat{X}_1) < 0$, then there is no vertex cover of size k, and if $V_1(\hat{X}_1) > 0$, then the subset of vertices $\hat{S} = \{v_l \in V : l \in \mathcal{L}(\hat{X}_1)\}$ must be a vertex cover of size k, given the result established above.

D The Inventory-Adjusted BoR Matching Policy

The inventory-adjusted policy. The inventory-adjusted BoR policy is a a roll-out version of (3), where the partitioning by ideal resource is defined with respect to the current inventory level, rather than the initial inventories. We formally define our roll-out policy through the binary indicators $u^t_{\ell,b}(X_t) \in \{0, 1\}$, which indicate whether a type b customer in period t is matched to resource ℓ, given that the current inventory levels are given by X_t. Specifically, we set

$$u^t_{\ell,b}(X_t) = \begin{cases} 1, & \text{if } b \in \mathcal{B}_\ell(X_t) \text{ and } r_{\ell,b} - \Delta\hat{V}^{\ell}_{t+1}(x_{\ell,t}; X_t) \geq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Letting $\mathcal{R}_t(X_t)$ denote the expected reward of this roll-out policy over periods $t, \dots, T$ when the period-t inventory vector is X_t, we have

$$\begin{aligned}\mathcal{R}_t(X_t) \quad = \quad & \sum_{b \in \mathcal{B}} p_{b,t} \cdot \left(\sum_{\ell \in \mathcal{L}(X_t)} u^t_{\ell,b}(X_t) \cdot \Big(r_{\ell,b} + \mathcal{R}_t(X_t - e_\ell) \Big) \right) \\ & \qquad + \left(1 - \sum_{b \in \mathcal{B}} \sum_{\ell \in \mathcal{L}(X_t)} p_{b,t} u^t_{\ell,b}(X_t) \right) \cdot \mathcal{R}_{t+1}(X_t)\end{aligned}$$

with base cases $\mathcal{R}_{t+1}(\cdot) = 0$. Implementing this roll-out policy in period t when the inventory vector is X_t, boils down to computing $\{u^t_{\ell,b}(X_t)\}_{\ell \in \mathcal{L}(x_t), b \in \mathcal{B}}$ via (17). These indicators can easily be derived after computing $\{\Delta V^\ell_t(x; X_t)\}_{\ell \in \mathcal{L}(X_t), x \in [x_{\ell,t}]}$ via (4). Noting that for any two inventory vectors $X, X' \in \mathbb{Z}^L_+$ that satisfy $\mathcal{L}(X) = \mathcal{L}(X')$, we have that $\hat{V}^\ell_t(x; X) = \hat{V}^\ell_t(x; X')$, it turns that we only need to recompute these decoupled value functions after each stock-out, of which there can be at most C.

The improved performance. The following lemma shows that the inventory-adjusted roll-out policy strictly improves upon the BoR policy of Section 2.1.

Lemma D.1. *For any period $t \in [T]$, starting inventory vector X_1, and period-t inventory level X_t that satisfies $x_{\ell,t} \leq x_{\ell,1}$ for each $\ell \in \mathcal{L}$, we have $\mathcal{R}_t(X_t) \geq \hat{V}_t(X_t; X_1)$, where $\hat{V}_t(X_t; X_1)$ is defined as in* (3).

Proof. We first establish the following intermediate claim, whose proof is presented at the end of the section.

Claim D.2. *For inventory vectors $X, X^< \in \mathbb{Z}^L_+$ that satisfy $\mathcal{L}(X^<) \subseteq \mathcal{L}(X)$, we have that $\hat{V}_t(X^<; X^<) \geq \hat{V}_t(X^<; X)$.*

Noting that Claim D.2 implies that $\hat{V}_t(X_t; X_t) \geq \hat{V}_t(X_t; X_1)$, we conclude the proof by establishing that $\mathcal{R}_t(X_t) \geq \hat{V}_t(X_t; X_t)$. To do so, we prove the more general result that $\mathcal{R}_t(X) \geq \hat{V}_t(X; X_t)$ for any inventory vector X that satisfies $\mathcal{L}(X) = \mathcal{L}(X_t)$. We prove this result via induction over t, noting that the base case of $t = T + 1$ trivially holds based on the terminal conditions of the two dynamic programs. Next, moving to the general case of $t \in [T]$, we note that $\hat{V}_t(X; X_t)$ can be expressed as

$$\begin{aligned}\hat{V}_t(X;X_t) &= \sum_{b\in\mathcal{B}} p_{b,t}\cdot\left(\sum_{\ell\in\mathcal{L}(X_t)} \hat{u}^t_{\ell,b}(x_\ell;X_t)\cdot\left(r_{\ell,b}+\hat{V}_t(X-e_\ell;X_t)\right)\right)\\ &\qquad + \left(1-\sum_{b\in\mathcal{B}}\sum_{\ell\in\mathcal{L}(X_t)} p_{b,t}\hat{u}^t_{\ell,b}(x_\ell;X_t)\right)\cdot\hat{V}_{t+1}(X;X_t),\end{aligned}$$

where

$$\hat{u}^t_{\ell,b}(x_\ell;X_t) = \begin{cases}1, & \text{if } b\in\mathcal{B}_\ell(X_t) \text{ and } r_{\ell,b}-\Delta\hat{V}^\ell_{t+1}(x_\ell;X_t)\ge 0\\ 0, & \text{otherwise.}\end{cases} \tag{18}$$

From here, we have

$$\begin{aligned}\hat{V}_t(X;X_t) &= \sum_{b\in\mathcal{B}} p_{b,t}\cdot\left(\sum_{\ell\in\mathcal{L}(X_t)} \hat{u}^t_{\ell,b}(x_\ell;X_t)\cdot\left(r_{\ell,b}+\hat{V}_{t+1}(X-e_\ell;X_t)\right)\right)\\ &\qquad + \left(1-\sum_{b\in\mathcal{B}}\sum_{\ell\in\mathcal{L}(X_t)} p_{b,t}\hat{u}^t_{\ell,b}(x_\ell;X_t)\right)\cdot\hat{V}_{t+1}(X;X_t)\\ &= \sum_{b\in\mathcal{B}} p_{b,t}\cdot\left(\sum_{\ell\in\mathcal{L}(X_t)} u^t_{\ell,b}(X)\cdot\left(r_{\ell,b}+\hat{V}_{t+1}(X-e_\ell;X_t)\right)\right)\\ &\qquad + \left(1-\sum_{b\in\mathcal{B}}\sum_{\ell\in\mathcal{L}(X_t)} p_{b,t}u^t_{\ell,b}(X)\right)\cdot\hat{V}_{t+1}(X;X_t)\\ &\le \sum_{b\in\mathcal{B}} p_{b,t}\cdot\left(\sum_{\ell\in\mathcal{L}(X_t)} u^t_{\ell,b}(X)\cdot\left(r_{\ell,b}+\hat{V}_{t+1}(X-e_\ell;X-e_\ell)\right)\right)\\ &\qquad + \left(1-\sum_{b\in\mathcal{B}}\sum_{\ell\in\mathcal{L}(X_t)} p_{b,t}u^t_{\ell,b}(X)\right)\cdot\hat{V}_{t+1}(X;X_t)\\ &\le \sum_{b\in\mathcal{B}} p_{b,t}\cdot\left(\sum_{\ell\in\mathcal{L}(X_t)} u^t_{\ell,b}(X)\cdot\left(r_{\ell,b}+\mathcal{R}_{t+1}(X-e_\ell)\right)\right)\\ &\qquad + \left(1-\sum_{b\in\mathcal{B}}\sum_{\ell\in\mathcal{L}(X_t)} p_{b,t}u^t_{\ell,b}(X)\right)\cdot\mathcal{R}_{t+1}(X)\\ &= \mathcal{R}_t(X).\end{aligned}$$

The first equality follows by comparing (17) and (18), and noting that in period t, both policies are identical as long as $\mathcal{L}(X)=\mathcal{L}(X_t)$. The first inequality uses Claim D.2, while the last inequality uses the induction hypothesis.

Proof of Claim D.2. We have that

$$
\begin{aligned}
\hat{V}_t(X^<; X^<) &= \sum_{\ell \in \mathcal{L}(X^<)} \hat{V}_t^\ell(x_{\ell,t}^<; X^<) \\
&\geq \sum_{\ell \in \mathcal{L}(X_t)} \hat{V}_t^\ell(x_{\ell,t}^<; X) \\
&= \sum_{\ell \in \mathcal{L}(X_1)} \hat{V}_t^\ell(x_{\ell,t}^<; X) \\
&= \hat{V}_t(X^<; X).
\end{aligned}
$$

The first inequality follows by Claim 2.3, which can be applied since $\mathcal{L}(X^<) \subseteq \mathcal{L}(X)$. The second equality follows by noting that if $\ell \in \mathcal{L}(X) \setminus \mathcal{L}(X^<)$, we must have $x_{\ell,t}^< = 0$ and hence $\hat{V}_t^\ell(x_{\ell,t}^<; X) = 0$. □

E The Costed ABI Trailer Problem

In this section, we present numerical experiments in which we carry out the approach outlined at the end of Section 4.4 to solve costed instances of the ABI Trailer Problem. Specifically, we apply our approach to the inner maximization problem of problem (10) for each $(W, C) \in [LT] \times [n]$. We benchmark the performance of this approach against the following upper bound for the costed variant of our problem: we solve a modified version of (Fluid-IP) in which (i) C and W are added as decision variables (C now is the number of distinct trailer types utilized and W is the total number of trailers loaded), (ii) the term $-Cost(W, C)$ is added to the objective, (iii) $W \geq T$ is added to the constraints, and (iv) x_l's are relaxed to be continuous variables. While this approach does not lead to concrete theoretical performance guarantees for problem (9), we find that its practical performance is near-optimal.

E.1 Choosing C alone

We first consider the case when $W = T$, and hence we can only vary the number of distinct traile types utilized. Consider $\text{Cost}(W, C) = k_1 \cdot C + \text{Cost}_w(W)$, e.g., ABI incurs a cost of k_1 for each unique trailer type that is stocked. Since W is not a decision variable here, we normalize $\text{Cost}_w(W) = 0$. Within our experiments, we set $k_1 = \delta \cdot Z^*(\infty, \infty)$ and vary $\delta \in \{0.0001, 0.001, 0.01\}$, so that the cost k_1 for each additional trailer type is some (potentially very small) fraction of the best case expected reward $Z^*(\infty, \infty)$. The results of our experiments are presented in Tables 4a and 4b.

Across all test cases, we observe that the largest optimality gap is 1.26%, and hence it is clear that this approach performs quite well. We also see the optimality gaps shrink as δ is increased,

					Our Approach		Fluid IP		
Carrier	T	W	δ	k_1	C^*	Exp. Rew.	C^*	UB	OPT GAP
GTGA	23	23	0.0001	1	12	9,976	11	10,030	0.54%
			0.001	10	2	9,919	5	9,976	0.58%
			0.01	100.5	1	9,775	2	9,784	0.09%
MTNF	22	22	0.0001	1.1	5	10,730	11	10,765	0.32%
			0.001	10.8	2	10,699	3	10,721	0.20%
			0.01	107.8	1	10,570	1	10,570	0.00%
PRIJ	33	33	0.0001	2.9	9	29,282	10	29,377	0.32%
			0.001	29.4	3	29,173	4	29,219	0.16%
			0.01	294.3	1	28,782	1	28,781	0.00%
WENX	34	34	0.0001	2.5	11	25,278	11	25,423	0.57%
			0.001	25.5	3	25,145	4	25,295	0.60%
			0.01	254.6	1	24,878	1	24,878	0.00%

(a) CRTV

					Our Approach		Fluid IP		
Carrier	T	W	δ	k_1	C^*	Exp. Rew.	C^*	UB	OPT GAP
TAMI	25	25	0.0001	0.7	8	6,965	12	7,013	0.68%
			0.001	7	2	6,945	3	6,974	0.41%
			0.01	70.3	1	6,850	1	6,850	0.00%
WENP	36	36	0.0001	1.7	6	16,796	11	16,857	0.36%
			0.001	16.9	2	16,731	4	16,770	0.23%
			0.01	168.9	1	16,509	1	16,509	0.00%
WERD	38	38	0.0001	0.9	3	8,685	12	8,771	0.97%
			0.001	8.8	3	8,662	6	8,720	0.67%
			0.01	87.8	1	8,492	2	8,503	0.14%
WERS	24	24	0.0001	0.7	4	7,112	10	7,203	1.26%
			0.001	7.2	2	7,087	5	7,162	1.06%
			0.01	72.1	1	6,988	2	7,009	0.30%

(b) FCL

Table 4: Optimality gaps of our approach when choosing C alone.

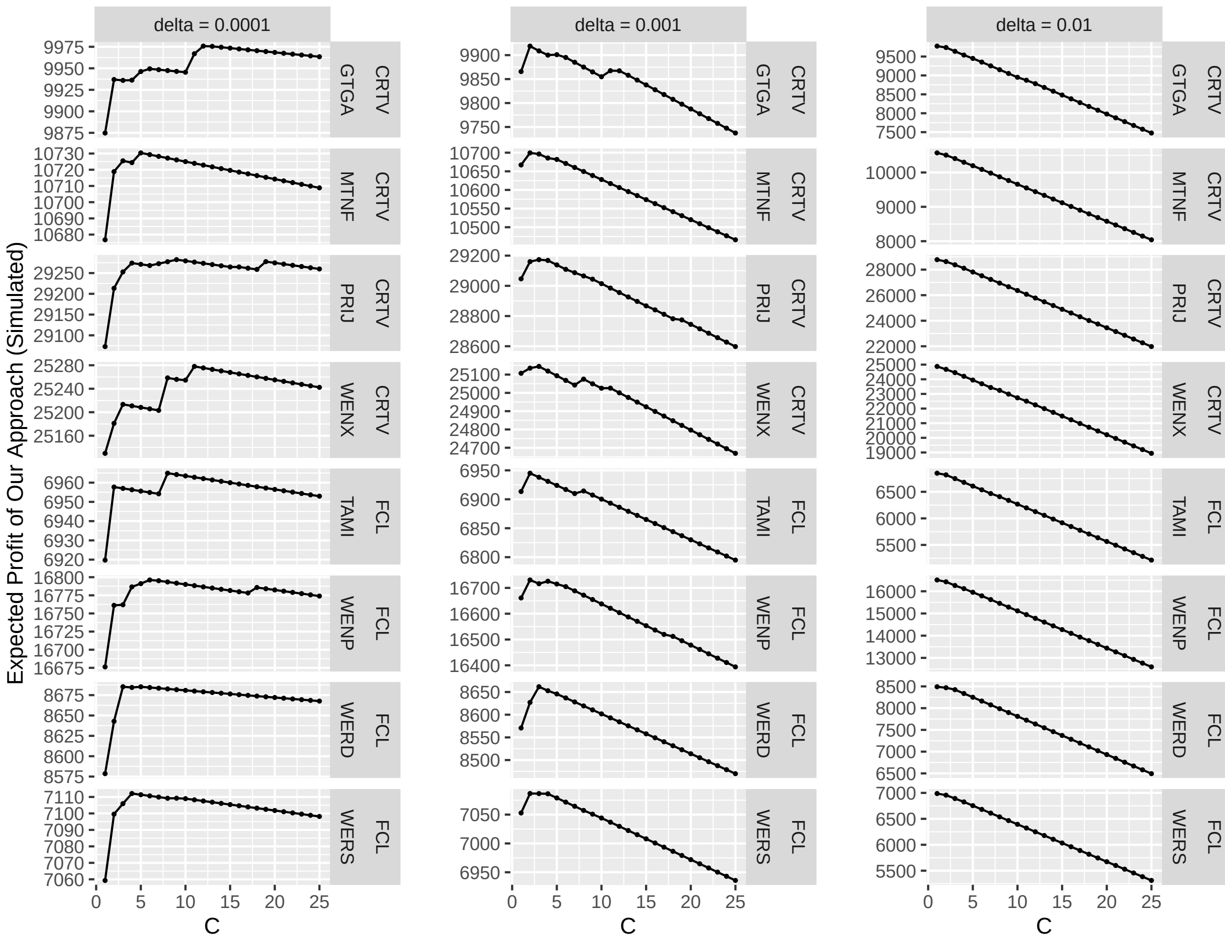

Figure 2: Illustration of Choosing C for the costed problem

which is likely a result of the fact that when $\delta = 0.01$ (and hence the operational cost is largest), it becomes optimal to stock only a single trailer type, and hence the problem is the "easiest". In other words, when the operational cost related to C becomes large enough so that utilizing a single trailer type is optimal, then the entire problem boils down to selecting the correct trailer type to stock, as the subsequent matching problem is trivial. Figure 2 provides a fine-grained illustration of how the expected profit changes as a function of C.

E.2 Choosing C and W

Next we test our approach for choosing C and W simultaneously. We consider a simple linear cost function given by $\text{Cost}(W, C) = k_1 \cdot C + k_2 \cdot W$, and vary $k_1, k_2 \in \{1, 10, 100, 200\}$. We provide our numerical result for carrier WENX (at warehouse CRTV) and WERS (at warehouse FCL) in Tables 5a and 5b. For the upper bound that is obtained from the revised Fluid-IP (see details at the beginning of this section), the optimal W is found to be equal to the number of truck arrivals T, which is not surprising and driven by the fact that the initial inventories are allowed to be fractional. We note that the observed optimality gaps remain very small, which further shows the efficacy of our approach when it is applied to a full costed version of our problem.

			Our Approach			Fluid IP			
Carrier	k_1	k_2	C^*	W^*	Exp. Rew.	C^*	W^*	UB	OPT GAP
WENX	1	1	21	44	25,378	17	34	25,410	0.126%
	1	10	14	39	25,013	17	34	25,104	0.361%
	1	100	11	34	21,893	17	34	22,044	0.683%
	1	200	11	34	18,493	17	34	18,644	0.807%
	10	1	7	38	25,283	5	34	25,332	0.192%
	10	10	5	37	24,945	5	34	25,026	0.324%
	10	100	8	34	21,797	5	34	21,966	0.768%
	10	200	8	34	18,397	5	34	18,566	0.909%
	100	1	2	35	25,000	2	34	25,070	0.279%
	100	10	1	34	24,692	2	34	24,764	0.292%
	100	100	1	34	21,632	2	34	21,704	0.334%
	100	200	1	34	18,232	2	34	18,304	0.396%
	200	1	1	34	24,898	1	34	24,898	0.001%
	200	10	1	34	24,592	1	34	24,592	0.001%
	200	100	1	34	21,532	1	34	21,532	0.001%
	200	200	1	34	18,132	1	34	18,132	0.002%

(a) CRTV

			Our Approach			Fluid IP			
Carrier	k_1	k_2	C^*	W^*	Exp. Rew.	C^*	W^*	UB	OPT GAP
WERS	1	1	19	47	16,811	18	36	16,831	0.121%
	1	10	18	37	16,450	18	36	16,507	0.348%
	1	100	6	36	13,199	18	36	13,267	0.512%
	1	200	6	36	9,599	18	36	9,667	0.702%
	10	1	4	38	16,729	5	36	16,763	0.204%
	10	10	4	36	16,393	5	36	16,439	0.281%
	10	100	4	36	13,153	5	36	13,199	0.349%
	10	200	4	36	9,553	5	36	9,599	0.480%
	100	1	1	36	16,542	2	36	16,543	0.003%
	100	10	1	36	16,218	2	36	16,219	0.003%
	100	100	1	36	12,978	2	36	12,979	0.004%
	100	200	1	36	9,378	2	36	9,379	0.005%
	200	1	1	36	16,442	1	36	16,442	0.000%
	200	10	1	36	16,118	1	36	16,118	0.000%
	200	100	1	36	12,878	1	36	12,878	0.000%
	200	200	1	36	9,278	1	36	9,278	0.000%

(b) FCL

Table 5: Optimality gaps of our approach when choosing both C and W.