Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Online Assortment Optimization with Reusable Resources

Xiao-Yue Gong, Vineet Goyal, Garud N. Iyengar, David Simchi-Levi, Rajan Udwani, Shuangyu Wang

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 68, No. 7, July 2022, pp. 4772–4785
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Online Assortment Optimization with Reusable Resources

Xiao-Yue Gong,[a] Vineet Goyal,[b] Garud N. Iyengar,[b] David Simchi-Levi,[c] Rajan Udwani,[d] Shuangyu Wang[c]

[a] Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; [b] Industrial Engineering and Operations Research, Columbia University, New York, New York 10027; [c] Institute for Data, Systems, and Society, Department of Civil and Environmental Engineering and Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; [d] Industrial Engineering and Operations Research, University of California Berkeley, Berkeley, California 94720

Contact: xygong@mit.edu (X-YG); vgoyal@ieor.columbia.edu, https://orcid.org/0000-0001-6719-3212 (VG); garud@ieor.columbia.edu, https://orcid.org/0000-0001-6546-4154 (GNI); dslevi@mit.edu (DS-L); rudwani@berkeley.edu, https://orcid.org/0000-0002-2112-4876 (RU); sw2756@columbia.edu (SW)

Received: January 22, 2020
Revised: December 25, 2020; March 30, 2021
Accepted: April 12, 2021
Published Online in Articles in Advance: November 11, 2021

https://doi.org/10.1287/mnsc.2021.4134

Abstract. We consider an online assortment optimization problem where we have n substitutable products with fixed reusable capacities $c_1, \ldots, c_n$. In each period t, a user with some preferences (potentially adversarially chosen) who offers a subset of products, S_t, from the set of available products arrives at the seller's platform. The user selects product $j \in S_t$ with probability given by the preference model and uses it for a random number of periods, $\tilde{t}_j$, that is distributed i.i.d. according to some distribution that depends only on j generating a revenue $r_j(\tilde{t}_j)$ for the seller. The goal of the seller is to find a policy that maximizes the expected cumulative revenue over a finite horizon T. Our main contribution is to show that a simple myopic policy (where we offer the myopically optimal assortment from the available products to each user) provides a good approximation for the problem. In particular, we show that the myopic policy is 1/2-competitive, that is, the expected cumulative revenue of the myopic policy is at least half the expected revenue of the optimal policy with full information about the sequence of user preference models and the distribution of random usage times of all the products. In contrast, the myopic policy does not require any information about future arrivals or the distribution of random usage times. The analysis is based on a coupling argument that allows us to bound the expected revenue of the optimal algorithm in terms of the expected revenue of the myopic policy. We also consider the setting where usage time distributions can depend on the type of each user and show that in this more general case there is no online algorithm with a nontrivial competitive ratio guarantee. Finally, we perform numerical experiments to compare the robustness and performance of myopic policy with other natural policies.

History: Accepted by Gabriel Weintraub, revenue management and analytics.
Funding: This work was supported by MIT-Accenture Alliance for Business Analytics, the MIT Data Science Lab, and Division of Civil, Mechanical, and Manufacturing Innovation [Grants 1351838 and 1636046].
Supplemental Material: Data are available at https://doi.org/10.1287/mnsc.2021.4134.

Keywords: assortment optimization online algorithms • reusable resources • coupling • competitive ratio

1. Introduction

Assortment optimization is an important problem that arises in a broad set of applications, including online advertising, recommendations, and e-retailing. In these applications, the goal of the decision-maker is to select a subset of products from the available universe to offer to the user to maximize the expected revenue or reward. For any given subset S of offered products, the selection of the user depends on the user's random preference over the set of products, including the no-purchase or exit option. We model this random selection using a *choice-model* that for any offer set S specifies the probability that the user selects product $j \in S \cup \{0\}$ (where 0 refers to the no-purchase or exit option). Several parametric choice models have been studied in the literature, including the multinomial logit (MNL) model (Luce 1959, McFadden 1973, Plackett 1975), the nested logit model (Williams 1977, McFadden 1978, Davis et al. 2014, Gallego and Topaloglu 2014), the Markov chain-based model Blanchet et al. (2016), and the mixture of multinomial logit model (McFadden and Train 2000) (see Train 2009, Kök et al. 2015, and Berbeglia et al. 2021 for a detailed overview of these models).

In this paper, we consider an online assortment problem where we are given n substitutable products with fixed capacities or inventories $c_1, \ldots, c_n$. Users with different choice models arrive sequentially. For each user, the seller offers a subset S of the available products satisfying certain constraints, and the user selects a random product $j \in S \cup \{0\}$ with probability given by his or her choice model and uses it for a random amount of time, $\tilde{t}_j$, and returns it to the platform, generating revenue $r_j(\tilde{t}_j)$ for the seller. The goal of the platform or the seller is to design a policy to offer assortments to the user so that the expected revenue is maximized.

This model fits the setting of classical online product allocation and revenue management. However, unlike traditional settings where the capacity or inventory of any product decreases permanently whenever the user selects that product, the products are "reusable" in our setting, meaning that upon allocation they come back into the inventory after some period of time and may be allocated several times over the planning horizon. Such a setting arises commonly in many applications, including cloud computing, physical storage, make-to-order service, and other sharing economy applications. For example, consider modern cloud platforms such as Amazon Web Services, Google Cloud, and Microsoft Azure. Among other services, these platforms commonly support large-scale data storage, a service that is widely used by online video platforms such as YouTube and Netflix. Given the large scale of these networks and huge volume of data, a given data file is stored in a subset of servers that are part of the cloud. Thus, a typical user request for data can be sent only to the subset of servers with the required data. Additionally, each server can concurrently serve only a limited number of requests while meeting the stringent low latency requirements on such platforms. Our online assortment model captures this setting as a special case where products correspond to servers in the cloud and starting inventory corresponds to capacity of servers. Customers correspond to user requests that arrive sequentially over time, and each customer must be irrevocably matched on arrival to at most one product out of a given subset that is revealed when the customer arrives. Interestingly, several recent works in the queuing theory study a similar setting with the objective of latency minimization on a bipartite networks, for instance, Mukherjee et al. (2018), Budhiraja et al. (2019), Weng et al. (2020), and Cruise et al. (2020). Our model ignores the queuing aspect and provides a complementary perspective from the point of view of maximizing the number of successful matches in a *loss system* (where the latency of accepted jobs is zero).

Settings similar to ours have previously been considered in the literature on online assortment planning for non-reusable products as well as reusable products. Golrezaei et al. (2014) considered the online assortment problem with fixed product capacities for the case of non-reusable products (i.e. inventory of a product decreases whenever any user selects that product). They gave an inventory balancing based algorithm that is $(1-1/e)$-competitive for adversarial arrivals in the limit of capacities going to infinity. For the case of all capacities being equal to one, their algorithm is $1/2$-competitive. Ma and Simchi-Levi (2020) considered a more general setting where the seller could make joint assortment and pricing decisions and obtain guarantees with adversarial customer arrivals for the case of non-reusable products.

Whereas the above results pertain to non-reusable resources, the most closely related result to our setting of reusable resources is in the work of Rusmevichientong et al. (2020). They considered this problem in the setting of stochastic customer arrivals where the distribution of user types is known in advance. Using this distributional knowledge, one can write the optimal algorithm, in this case as a dynamic program (DP), but this DP suffers from the curse of dimensionality. Rusmevichientong et al. (2020) gave an algorithm based on approximate dynamic programming and showed that it is a $1/2$-approximation to the optimal DP for the problem. Earlier, Dickerson et al. (2018) considered the problem of online matching (instead of assortments) when resources are reusable and the distribution of customer types is known. They proposed a simulation- and LP-based approach and showed that it is $1/2$-competitive against an offline LP benchmark (which is a stronger guarantee).

In contrast, we assume no advance knowledge of the user type distribution and consider an adversarial model for the sequence of user types. Recall that user type refers to the choice model of the user, which is revealed to the platform when the user arrives. Let ϕ^z be the choice model for user type z and $\phi^z(i,S)$ the probability that user type z selects product i given assortment S. We make the following assumptions about choice probabilities and usage time distributions.

Assumption 1. *For any user type z, assortment $S \subseteq T \in \mathcal{S}$ and $i \in S$, we have $\phi^z(i,S) \geq \phi^z(i,T)$.*

This is a mild assumption and without much loss of generality. In fact, all random utility-based choice models, including multinomial logit (MNL), nested logit, and mixture of MNLs, satisfy Assumption 1.

Assumption 2. *For every product j, usage time distribution depends only on j and not on the user type.*

For settings where this assumption does not hold, we show that it is impossible to obtain any constant factor competitive algorithm for adversarial arrivals. In general, Assumption 2 is reasonable in various settings where the choice of the product depends on the user type, whereas the usage time depends only on the product. In the setting of cloud computing, this translates to the assumption that the time taken to fulfill a user request depends primarily on the characteristics (processing power, memory) of the server that a request is assigned to. Although imperfect, this assumption is commonly employed in related literature (Mukherjee et al. 2018, Budhiraja et al. 2019, Cruise et al. 2020, Weng et al. 2020). As another example, consider a make-to-order setting where user selects a product from the offered assortment. Once the user makes the selection, a dedicated machine (resource) makes the product for the user. In such a setting, the busy time of the machine (resource) depends only on the product. Make-to-order settings also involve non-reusable resources in the form of raw materials. However, in a typical setting the machines are the bottleneck resources, because they are often much more expensive than the raw materials. It is also worth noting that in many scenarios, making a product may require the use of several machines either in parallel or in a given sequence. This presents new challenges that are not captured in the model we consider here and could be an interesting direction for future work.

The revenue to the seller, $\boldsymbol{r}_j(\tilde{t}_j)$ when product j is used for some random time, $\tilde{t}_j$, could be a general function of the usage time. In particular, we can model fixed revenue for every use as well as revenue that is an affine function of usage time with fixed component and per-unit usage time component. Let r_j denote the expected revenue of product j where the expectation is taken over the random usage time of product j, that is,

$$r_j = \mathbb{E}_{t_j \sim F_j}[\boldsymbol{r}_j(t_j)],$$

where F_j is the cdf of usage time distribution of product j.

1.1. Mathematical Formulation

We are given n substitutable products with reusable capacities $c_1, \ldots, c_n \in \mathbb{N}$. Each product i has a price $r_i \in \mathbb{R}^+$. In each period t over a horizon of length T, a customer, denoted by customer t, arrives to our platform. The customer's choice model ϕ^t becomes known to us upon arrival. We want to offer a subset of products S_t to this customer from the set of available products at time t. The expected revenue of any assortment S is $\sum_{i \in S} r_i \phi^t(i, S)$.

For any given subset S_t, a user at time t selects a product or no purchase according to the user's choice model ϕ^t. If user t selects a product j, he or she uses it for a random number of periods, $\tilde{t}_j$, that is distributed i.i.d. according to some distribution that depends only on product j. This purchase generates a revenue $r_j(\tilde{t}_j)$ for us before the product comes back into the inventory. Full revenue is collected regardless of the end of the horizon. At the end of the horizon no further customers arrive, but the revenue is still collected from resources in use until they are returned.

We assume Assumptions 1 and 2. The arrival sequence of customers and their respective preferences can be adversarially chosen, but the adversary is oblivious. In other words, we make no assumptions (distributional or otherwise) on the arrival sequence and allow it to be completely arbitrary (periodic or aperiodic/continuous). The goal of the seller is to find a policy that maximizes the expected cumulative revenue over a finite horizon T,

$$\max \mathbb{E}\left[\sum_{t=1}^{T} \sum_{j \in S_t} r_j \phi^t(j, S_t)\right],$$

where the expectation is over product choices and realizations of usage times.

We take the perspective of competitive analysis on our policy, where we measure the worst-case performance of a policy against a clairvoyant optimal policy that is computed with the complete knowledge of arrival epochs, the choice models of all customers, and all the usage time distributions. Note that OPT does not know the realizations of product choices and usage times beforehand. We compare the worst-case expected revenue collected by our policy with the optimal expected revenue collected by the optimal policy, where the expectation is again over product choices and realizations of usage times.

1.2. Our Contributions

Our main contribution is to show that a myopic policy provides a good approximation for the online assortment optimization problem with reusable products.

1.3. Myopic Policy

For each user, the myopic policy offers an assortment $S \in \mathcal{S}$ from the set of available products that maximizes the expected revenue from that user. More specifically, suppose user at time t has type z_t and let $\mathcal{I}_t$ be the set of products available to the myopic policy at time t. Then the myopic policy offers assortment S_t, where

$$S_t \in \operatorname{argmax}\left\{\sum_{i \in S} r_i \cdot \phi^{z_t}(i, S) \,\middle|\, S \subseteq \mathcal{I}_t, S \in \mathcal{S}\right\},$$

where recall that ϕ^z is the choice model for user type z, $\phi^z(i, S)$ is the probability that user type z selects product i given assortment S, and r_t is the expected revenue when product i is selected, where the expectation is taken over the usage time of product i. Recall that we assume that the usage time distribution depends only on the product and is not dependent on user type. Therefore, the myopic policy only needs the expected revenue r_j from product j if it is selected and does not need any further information about the usage time distribution. Furthermore, the optimal set S_t can be found using any black box algorithm for static assortment optimization.

We show that this myopic policy is 1/2-competitive. In other words, the expected revenue of the myopic policy is at least 1/2 times the expected revenue of an optimal policy that has full information about the sequence of user types and product usage distributions (albeit not the choice realizations and the realization of usage times). We refer to this as the *clairvoyant benchmark*. More generally, when the (possibly constrained) static assortment optimization problem at each stage can only be solved up to within an α factor of the optimal, our myopic policy is $\alpha/(1+\alpha)$-competitive. We would also like to remark that even for the case of non-reusable items, which is a special case of our setting, there are instances where the myopic policy achieves exactly 1/2 the value of clairvoyant (for examples, see Karp et al. 1990 and Golrezaei et al. 2014).

1.4. Impossibility Result for User Type-Dependent Usage Distributions

We also show that if the usage time distribution depends on the user type, then there is no online algorithm that can obtain a constant factor competitive ratio as compared with our clairvoyant benchmark for the case of adversarial arrivals. This result holds even in the large-capacity case. We would like to note that Rusmevichientong et al. (2020) considered the case where usage time distributions can depend on the user type. However, they considered the setting of stochastic arrivals with known distribution of user types and give a 1/2-approximation compared with the optimal dynamic programming solution as opposed to the clairvoyant benchmark.

1.5. Challenges and New Techniques

We would like to note that even with full-information about the sequence of users and the usage time distribution, computing an optimal policy is intractable due to the curse of dimensionality. Golrezaei et al. (2014) used an LP-based upper bound as a benchmark for the case of non-reusable products. One of the challenges in extending the results to the case of reusable products is the lack of a good LP-based upper bound. The LP formulation in Golrezaei et al. (2014) has an unbounded gap because it does not account for reusability and, therefore, is not a useful benchmark for the problem. On the other hand, LP upper bounds of the form proposed in Dickerson et al. (2018) naturally capture reusability; however, it is not immediately clear how to perform a primal-dual type analysis with this more involved LP. Part of the reason is that, because of reusability, the remaining capacity of products is nonmonotonic (capacity decreases when product is used but increases when used units return). But more importantly, this nonmontonicity is inherently stochastic due to the uncertainty in usage durations, and this makes it nontrivial to dual-fit in the way of Devanur et al. (2013) and Golrezaei et al. (2014).

In order to prove the competitive ratio bound, we design a novel *queue-based coupling* that allows us to upper bound the expected revenue of an optimal algorithm with full information in terms of the expected revenue of the myopic policy for any usage time distributions.

1.6. Further Work Since This Paper

Ever since a version of this paper appeared online, there have been several developments in these settings, especially in the large-capacity regime. In case of adversarial arrivals, Feng et al. (2019) showed that the inventory balancing algorithm of Golrezaei et al. (2014) is $(1-1/e)^2 \approx 0.4$-competitive for large capacity. For the special case of deterministic/fixed usage durations, they showed the best possible performance guarantee of $(1-1/e)$ in the large-capacity regime. In contrast, Goyal et al. (2021) (merging two earlier papers, Goyal et al. 2020a and Goyal et al. 2020b) demonstrated that the general case of stochastic usage durations is fundamentally different, and inventory balancing may not be sensitive enough to address reusability even for simple two-point usage distributions. First, they proposed a new ranking-based allocation scheme and showed that it achieves the best possible guarantee of $(1-1/e)$ for large capacities when the usage distributions are IFR (roughly speaking). Building on this insight, they develop a fluid-guided algorithm that is $(1-1/e)$-competitive for arbitrary usage distributions when the capacities are large. To analyze these algorithms, they introduced a new LP-free analysis approach inspired by the primal-dual analysis of Devanur et al. (2013) and its path-based generalization in Goyal and Udwani (2020). Beating the performance of the myopic algorithm without the large-capacity assumption remains open.

In the stochastic arrival setting, Feng et al. (2019) showed a competitive ratio of 1/2 against a stronger LP benchmark and with a different simulation-based policy (recall that Dickerson et al. (2018) showed such a result for the special case of online matching of reusable resources). In concurrent work, Baek and Ma (2019) gave a 1/2-competitive policy against the LP benchmark more generally for network revenue management with reusable resources. Finally, Feng et al. (2020) showed a near optimal result in the stochastic arrival case for large capacities.

1.7. Related Work

There is a considerable amount of literature on dynamic assortment optimization problems with non-reusable products, starting with Bernstein et al. (2015), which studied the problem of dynamic assortment optimization for a stochastic arrival model where users choose according to a multinomial logit choice model (Gallego et al. 2004, Talluri and Van Ryzin 2004, Liu and van Ryzin 2008, Topaloglu 2013) and the user type is drawn i.i.d. from a stationary distribution. Chan and Farias (2009) considered a stochastic depletion framework for nonstationary environments that includes the assortment-planning problem under random arrivals. They gave a 1/2-competitive myopic policy for this general framework. More recently, Stein et al. (2018) and Wang et al. (2018) considered other closely related models for online product allocation with stochastic arrivals. We refer the reader to Golrezaei et al. (2014) for a more detailed review.

For revenue management with reusable products and random usage times, Levi and Radovanović (2010) first studied a product-independent demand model where the users do not exhibit any choice behavior and the goal is to design a policy to maximize the average revenue in an infinite horizon setting. Owen and Simchi-Levi (2018) extended this model to include user choices and also studied the infinite horizon setting. Chen et al. (2017) considered a related problem of control admission for a system with multiple units of a single product that can be reserved in advance for time intervals determined by users arriving according to a multiclass Poisson process.

Product allocation problems also closely relate to online matching problems and often generalize the classical online bipartite matching problem of Karp et al. (1990). In this seminal work, they showed that matching arriving users (the unknown vertices) based on a random ranking over all products (vertices on the known side of the graph) gives the best possible competitive guarantee of $(1-1/e)$. We refer the interested reader to Mehta (2013) for a more detailed review of work on online matching and its variants/generalizations.

1.8. Outline

The rest of the paper is organized as follows. In Section 2, we present the competitive ratio analysis for general usage time distributions. In Section 3, we present a family of examples to show that no online algorithm can have a constant competitive ratio for adversarial arrivals if the usage time distributions depend on the user type. In Section 4, we include some promising results comparing our myopic policy with other known policies on a synthetic testbench. Finally, in Section 5, we summarize our results and discuss some directions for future work.

2. Competitive Ratio of Myopic Policy

In this section, we show that the myopic policy is 1/2-competitive for general usage time distributions and general revenue functions (as functions of usage times). Before proceeding, we discuss useful notation and introduce an important simplification.

2.1. Notation

Let $\mathcal{N}=\{1,\ldots,n\}$ be the set of products available to the platform with capacities $c_1,\ldots,c_n$. Let us refer to the myopic policy as ALG and the clairvoyant optimal as OPT. Recall that we let $\mathcal{S}$ denote the set of feasible assortments. Let $\mathcal{Z}$ denote the set of user types. For any $z\in\mathcal{Z}$, $S\subseteq\mathcal{N}$, $i\in S\cup\{0\}$, let $\phi^z(i,S)$ denote the probability that user type z selects i when offered assortment S. The choice probabilities satisfy Assumption 1. Also, recall that r_j is the expected revenue to the seller when j is selected by any user and F_j is the cdf of usage duration for product j. $R(S,z_t)=\sum_{i\in S} r_t\cdot\phi^{z_t}(i,S)$ is the expected revenue of assortment S for user type z_t. Let ω denote the sample path that specifies the random preference realizations of all users $z_1,\ldots,z_T$ and the random usage times. Let $\mathcal{I}_t(\omega)$ denote the set of available products in ALG at time t on sample path ω. Also, let $S_t(\omega)$ ($S^*_t(\omega)$, respectively) denote the assortment offered by ALG (OPT, respectively) at time t to user type z_t. Here, recall that $S_t(\omega)=\operatorname{argmax}_{S\in\mathcal{I}_t(\omega)} R(S,z_t)$. Let $j_t(\omega)\in S_t(\omega)\cup\{0\}$ be the product selected by user z_t in ALG, and let $j^*_t(\omega)\in S^*_t(\omega)\cup\{0\}$ be the product selected by user z_t in OPT at time t on sample path ω. Note that product 0 refers to the do-nothing or exit option.

2.2. From Arbitrary to Unit Inventory

We now argue that one can safely assume that $c_j=1$ for all j, since a guarantee for the case of unit inventory leads to a stronger result that generalizes to the case of arbitrary inventories. This allows us to perform a simpler and crisper analysis without loss of generality (w.l.o.g.). We note that such a property is commonly used in the online matching literature for non-reusable resources (Mehta et al. 2013), and we show this also for our assortment setting with reusable resources.

Given a setting with arbitrary inventories, $\{c_j\}_{j\in\mathcal{N}}$. Consider a unit inventory setting where for each j we have c_j identical products (with the same usage distribution as the original product) instead of c_j units of product j in the original instance. We refer to this as the *unit setting*. In the unit setting, we index the products as (j,k_j), where $k_j\in[c_j]$ for every $j\in\mathcal{N}$. For any given assortment S_u in the unit setting, we let S denote the set of products in the original instance with $j\in S$ if there exists some $k_j\in[c_j]$ such that $(j,k_j)\in S_u$. Now, given an arrival with choice model ϕ on the original product space, define the following choice function ϕ_u in the new space of products:

$$\phi_u((j,k_j),S_u)=\begin{cases}\phi(j,S) & \text{if for every } k<k_j,\ (j,k)\notin S_u,\\ 0 & \text{otherwise.}\end{cases}$$

Furthermore, we couple the customer choice between the two settings so that when the customer chooses product $j\in S$ in the original instance, the unique product $(j,k_j)\in S_u$ is such that $\phi_u((j,k_j),S_u)=\phi(j,S)$ is chosen in the unit setting and vice versa.

Note that if choice model ϕ satisfies Assumption 1, then so does the unit setting choice model ϕ_u. Given an arrival sequence in the original setting, we construct an arrival sequence for the unit setting using this transformation. The following proposition shows that the expected revenue of OPT (and ALG) does not change when we perform this transformation.

Proposition 1. *Given an instance of the problem with arbitrary inventories $\{c_i\}_{i\in\mathcal{N}}$, the expected total revenue of OPT and ALG remains unchanged in the transformed instance with unit inventories.*

Proof. Consider an arbitrary arrival sequence in the original instance and its equivalent sequence (as given by the transformation) in the unit setting. Given an optimal algorithm (OPT) for the original instance, we construct an algorithm for the unit setting with the same expected revenue. For every product j, whenever OPT includes j in an assortment for the original instance, we include exactly one available product (j,k_j) for an arbitrary $k_j\in[c_j]$ in the unit setting. This defines a policy for the unit setting, but to make this definition meaningful we need to ensure that whenever j is available in the original instance, some product (j,k_j) is available in the unit instance. This is true at the first arrival. Inductively, following the defined policy while using the same realization of usage times for both settings and coupling the customer choice as described earlier, we have that if product j is available in the original instance at arrival t, then we are guaranteed some k_j such that (j,k_j) is available in the unit setting at t. Therefore, we have a well-defined policy for the unit setting with expected revenue exactly as much as OPT in the original setting.

The reverse is also true. Given the optimal algorithm for the unit setting, whenever a product (j,k_j) is included in the assortment, we include product j in the assortment for the original instance. Once again, because of the coupling between the choice models and using the same realizations of usage durations, whenever a product (j,k_j) is available in the unit setting, at least one unit of j is available in the original setting. This leads to a policy for the original setting with the same expected revenue as the optimal policy for the unit setting. Therefore, the optimal value of clairvoyant is the same in both instances. In fact, using the same argument, the expected revenue of ALG is also identical in the two settings. □

In the rest of this section, we will show that on every unit inventory instance the expected revenue of ALG is at least half of OPT. The above proposition then gives us our general result for arbitrary inventories. Note that, although the number of products in the unit setting can be much larger, this is only for the purpose of analysis and has no impact on the actual run time of the myopic policy (which is on the original space of products and choice models).

Theorem 1. *Suppose for every product j, the usage time is distributed according to a distribution that only depends on the product j itself. Then, for any sequence of user types $z_1,\ldots,z_T$, the expected cumulative revenue of the myopic policy is at least $1/2$ times the expected cumulative revenue of the clairvoyant optimal that knows the full sequence, that is,*

$$\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega),z_t)\right]\geq\frac{1}{2}\cdot\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega),z_t)\right].$$

Proof. The proof proceeds by decomposing the revenue $\mathbb{E}_\omega[R(S_t^*(\omega),z_t)]$ of OPT into two parts, one corresponding to the products in $S_t^*(\omega)$ that are available in ALG at time t and the other corresponding to products in $S_t^*(\omega)$ that are unavailable in ALG at time t. We upper bound the former by the revenue $\mathbb{E}_\omega[R(S_t(\omega),z_t)]$ of ALG for the same arrival. The key challenge is to bound the total expected revenue in OPT from products that are unavailable in ALG at the time they are offered in OPT. We do this subsequently in Lemma 2. Formally,

$$\begin{aligned}\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega),z_t)\right]&=\mathbb{E}_\omega\left[\sum_{t=1}^{T}\sum_{j\in S_t^*(\omega)}r_j\cdot\phi^{z_t}(j,S_t^*(\omega))\right]\\
&=\mathbb{E}_\omega\left[\sum_{t=1}^{T}\left(\sum_{j\in S_t^*(\omega)\cap\mathcal{I}_t(\omega)}r_j\cdot\phi^{z_t}(j,S_t^*(\omega))\right.\right.\\
&\qquad\left.\left.+\sum_{j\in S_t^*(\omega)\backslash\mathcal{I}_t(\omega)}r_j\cdot\phi^{z_t}(j,S_t^*(\omega))\right)\right]\\
&\leq\mathbb{E}_\omega\left[\sum_{t=1}^{T}\sum_{j\in S_t^*(\omega)\cap\mathcal{I}_t(\omega)}r_j\cdot\phi^{z_t}(j,S_t^*(\omega))\right]\\
&\qquad+\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega),z_t)\right]\\
&\leq\mathbb{E}_\omega\left[\sum_{t=1}^{T}\sum_{j\in S_t^*(\omega)\cap\mathcal{I}_t(\omega)}r_j\cdot\phi^{z_t}(j,S_t^*(\omega)\cap\mathcal{I}_t(\omega))\right]\\
&\qquad+\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega),z_t)\right]\\
&\leq2\cdot\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega),z_t)\right],\end{aligned}$$

The first inequality follows from the claim $\mathbb{E}_\omega[\sum_{t=1}^T \sum_{j\in S_t^*(\omega)\setminus\mathcal{I}_t(\omega)} r_j \cdot \phi^{z_t}(j, S_t^*(\omega))] \le \mathbb{E}_\omega[\sum_{t=1}^T R(S_t(\omega), z_t)]$, which we show in Lemma 2. The second inequality follows from Assumption 1. The final inequality follows from the definition of $S_t(\omega)$, which is a subset of $\mathcal{I}_t(\omega)$ that maximizes the single period revenue for user type z_t. □

2.3. Queue-Coupling Technique

In order to bound the total expected revenue in OPT from products that are unavailable in ALG at the time they are offered in OPT, we introduce a new coupling between the usage times in ALG and OPT. In particular, we introduce *coupling queues* to specify the coupling of usage times between sample paths in ALG and OPT. For each product j, we maintain a queue, $\mathcal{Q}_j$. Initially, $\mathcal{Q}_j$ is empty.

Whenever product j is selected in ALG by any user, we generate an i.i.d. sample from the usage time distribution F_j and insert the sample at the rear of the queue, $\mathcal{Q}_j$.

Whenever product j is selected in OPT by any user, we get the first element of queue, $\mathcal{Q}_j$, and use it as a usage time sample for product j in OPT (we also remove this element from the queue). So, we use the samples in $\mathcal{Q}_j$ in a FIFO order. This couples the usage distributions of ALG and OPT. In case $\mathcal{Q}_j$ is empty, we generate an i.i.d. sample from F_j.

Lemma 1. *For any time $t = 1, \ldots, T$ and any product $j = 1, \ldots, n$, whenever a user selects product j in OPT, the usage time distribution given by the above coupling is i.i.d. according to the usage time distribution for product j.*

Proof. The interesting case is when $\mathcal{Q}_j$ is not empty. Then, the sample that is used by OPT and removed from the queue, denoted by $\tilde{L}_j$, was originally picked independently from all previous samples and added to the queue unconditionally. Any other samples that might have been added to the queue subsequent to adding $\tilde{L}_j$ do not affect $\tilde{L}_j$. All samples in a queue have the same probability of being the front of the queue. Therefore, the samples obtained from the queues by selection of the product in OPT are i.i.d. according to F_j. □

We are now ready to bound the total expected revenue in OPT from products that are unavailable in ALG at the time they are offered in OPT. In particular, we have the following lemma.

Lemma 2. *For any usage time distributions, and sequence of user types $z_1, \ldots, z_T$,*

$$\mathbb{E}_\omega\left[\sum_{t=1}^T \sum_{j\in S_t^*(\omega)\setminus\mathcal{I}_t(\omega)} r_j \cdot \phi^{z_t}(j, S_t^*(\omega))\right] \le \mathbb{E}_\omega\left[\sum_{t=1}^T R(S_t(\omega), z_t)\right].$$

Proof. Consider any ω such that $j_t^*(\omega) \in S_t^*(\omega)\setminus\mathcal{I}_t(\omega)$. Let us refer to $j_t^*(\omega)$ as j_t^* for brevity. At time t, j_t^* is not available in ALG on sample path ω. Therefore, it is in use at time t and must have been selected in ALG at some previous time period, say $t-\tau$, for some $\tau \ge 1$. Let $\tilde{L}$ be the random usage time that ALG sampled for j_t^* at time $(t-\tau)$ and inserted in the queue $\mathcal{Q}_{j_t^*}$ corresponding to product j_t^*. Because j_t^* is still in use by ALG by our coupling, we get $\tilde{L} \ge \tau$. Using this and the fact that OPT is able to select j_t^* at time t, we have that the sample $\tilde{L}$ must exist on the queue up to time t (but may be popped at t). Therefore, $\mathcal{Q}_{j_t^*}$ is nonempty before user arrives at t.

Hence, when OPT selects j_t^* at time t, we get a sample from $\mathcal{Q}_{j_t^*}$. Suppose the sample used by OPT was generated for ALG at time $t' \le t-\tau$. We charge the revenue earned by OPT for this selection to the revenue earned by ALG for using j_t^* at time t'. Observe that the charging is unique since each sample on the queue is used at most once by OPT, and we only charge to ALG when the corresponding sample is used by OPT. Therefore,

$$\begin{aligned}\mathbb{E}_\omega\left[\sum_{t=1}^T \sum_{j\in S_t^*(\omega)\setminus\mathcal{I}_t(\omega)} r_j \cdot \mathbb{1}(j = j_t^*(\omega))\right] &\le \mathbb{E}_\omega\left[\sum_{t=1}^T r_{j_t(\omega)}\right] \\ &= \mathbb{E}_\omega\left[\sum_{t=1}^T R(S_t(\omega), z_t)\right]. \end{aligned} \tag{1}$$

We would also like to note that the revenue from a product can now even depend on the usage time duration of the product. Simplifying as before, we get

$$\begin{aligned}&\mathbb{E}_\omega\left[\sum_{t=1}^T \sum_{j\in S_t^*(\omega)\setminus\mathcal{I}_t(\omega)} r_j \cdot \phi^{z_t}(j, S_t^*(\omega))\right] \\ &\quad\le \mathbb{E}_\omega\left[\sum_{t=1}^T \sum_{j\in S_t^*(\omega)\setminus\mathcal{I}_t(\omega)} r_j \cdot \phi^{z_t}(j, S_t^*(\omega)\setminus\mathcal{I}_t(\omega))\right] \\ &\quad= \mathbb{E}_\omega\left[\sum_{t=1}^T \sum_{j\in S_t^*(\omega)\setminus\mathcal{I}_t(\omega)} r_j \cdot \mathbb{1}(j = j_t^*(\omega))\right] \\ &\quad\le \mathbb{E}_\omega\left[\sum_{t=1}^T R(S_t(\omega), z_t)\right],\end{aligned}$$

where the first inequality follows from Assumption 1 and the last inequality follows from (1). □

2.4. Generalizations and Extensions

The following results follow as a direct consequence of our main result.

Corollary 1. *Given an α-approximation algorithm for solving the (possibly constrained) static assortment optimization problem at each stage, our myopic policy is $\alpha/(1+\alpha)$-competitive.*

Proof. Recall the proof of Theorem 1, where we split the revenue of OPT into two parts, one corresponding to the products in $S_t^*(\omega)$ that are available in ALG at time t and the other corresponding to products in $S_t^*(\omega)$ that are unavailable in ALG at time t. The latter term is still bounded as before since Lemma 2 holds even if the static assortment optimization problem at each stage can be solved only approximately. However, the former is now bounded as follows,

$$\mathbb{E}_\omega\left[\sum_{t=1}^{T}\sum_{j\in S_t^*(\omega)\cap\mathcal{I}_t(\omega)} r_j\cdot\phi^{z_t}(j,S_t^*(\omega)\cap\mathcal{I}_t(\omega))\right] \le \frac{1}{\alpha}\cdot\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega),z_t)\right].$$

This results in a competitive ratio $\alpha/(1+\alpha)$. □

The following corollary addresses situations where the revenue from each product is nonstationary and varies across time.

Corollary 2. *Letting γ denote the ratio of minimum to maximum revenue of any product across all arrivals, the myopic policy is $\frac{\gamma}{2}$-competitive.*

The proof follows directly by considering the worst-case scenario where the revenue collected by our myopic policy is always with maximum markdown, whereas none of the revenue collected by OPT is discounted. As an example of application of the corollary, if the maximum markdown on products over the entire planning horizon is 10%, then $\gamma = 0.9$, and the myopic policy guarantees a total revenue at least 0.45 times that of the clairvoyant.

2.5. Tighter Result Using Booking Limits

This result above is arbitrarily bad when $\gamma \to 0$, that is, when products may be sold to customers at a steep discount. In this case, we can do better by setting random booking limits. This idea is inspired by a similar notion in Ball and Queyranne (2009) for a setting with non-reusable resources. It generalizes naturally to the reusable case and leads to a worst-case guarantee of $\frac{1}{4}$.

Formally, for any product j, let r_j denote the normal price of the item, and let r_{jd} denote the discounted price. We treat each product as two separate products and assume that the choice model for every arriving customer dictates (possibly) different probabilities based on prices. Recall that $\gamma = \min_j \frac{r_{jd}}{r_j}$. The new myopic policy works as described next. In the beginning, we randomly decide whether to offer products at a discount. Specifically, with a probability of 0.5 we consider both the discounted and normal versions of all products when finding optimal assortment for customers, and with the remaining probability of 0.5 we do not include the discounted version of products in any assortment. This protects against the possibility of selling products at steep discounts when future arrivals would have chosen the same product at higher prices. After randomly pruning the discounted products at the start of the planning horizon, to each customer we offer the revenue maximizing assortment.

Lemma 3. *The myopic policy with random booking limit is $\frac{1}{4}$-competitive even for $\gamma \to 0$.*

Proof. On each sample path ω, let $\mathcal{I}_t(\omega,d)$ denote the set of products available at t in ALG when discounted products are included. We write $\mathcal{I}_t(\omega)$ to denote the set of available products when discounted products are blocked. Let $S_t^*(\omega)$ and $S_t^*(\omega,d)$ denote the subset of normal price products and the subset of discounted products offered to arrival t in OPT. Let $S_t(\omega)$ denote the assortment offered to t in ALG when discounted products are excluded, and let $S_t(\omega,d)$ denote the overall assortment when discounted products are included in ALG. Finally, let ALG_d denote the expected revenue of ALG given that discounted products are included, and similarly, ALG_n denotes the expected revenue given that discounted products are excluded. Overall, $\text{ALG} = 0.5(\text{ALG}_d + \text{ALG}_n)$. Now,

$$\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega)\cup S_t^*(\omega,d),z_t)\right] \le \mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega),z_t)\right] + \mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega,d),z_t)\right].$$

Applying the decomposition used in proving Theorem 1 and using the coupling argument from Lemma 2, it follows that

$$\begin{aligned}\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega),z_t)\right] &\le \mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega)\cap\mathcal{I}_t(\omega),z_t)\right] \\ &\quad + \mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega),z_t)\right] \\ &= 2\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega),z_t)\right] = 2\text{ALG}_n.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega,d)),z_t)\right] &\le \mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t^*(\omega,d)\cap\mathcal{I}_t(\omega,d),z_t)\right] \\ &\quad + \mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega,d),z_t)\right] \\ &\le 2\mathbb{E}_\omega\left[\sum_{t=1}^{T}R(S_t(\omega,d),z_t)\right] = 2\text{ALG}_d.\end{aligned}$$

Combining the inequalities, we have $\text{OPT} \le 4\text{ALG}$, as desired. □

3. User Type-Dependent Usage Times: Family of Bad Examples

We now consider the case where the product usage time distributions could depend on the user type and show that there is no online algorithm with a constant competitive ratio in the adversarial arrival model. The result holds even in the high-capacity regime, where the capacities of all products go to $+\infty$. It will suffice for us to consider a single product. For a user arriving at time t, let d_t denote the random usage duration. Even for this special case with a single reusable product we have the following upper bound on the competitive ratio of any online algorithm.

Theorem 2. *For online matching with a single reusable product and arbitrary product capacity, if the random usage durations depend on the user, no online (randomized) algorithm can have a competitive ratio better than* $O\left(\frac{\log T}{T}\right)$, *where T is the number of users.*

Before we prove the above theorem, consider first the special case of algorithms that always match a user to some available product if such a matching is possible. Suppose we have a single unit of a single product with reward 1 and the following arrival sequences:

• *Sequence A*: A single user with usage time duration ∞ (never returns the product).

• *Sequence B*: A user with usage duration ∞, followed by T users that return the product right away, that is, $\mathbb{P}(d_t = 0) = 1$ for all $t \in \{2, \ldots, T+1\}$.

In order to be competitive on *sequence A*, the algorithm must match the arrival with the only available product. Consequently, even on *sequence B*, the algorithm will match the product to the first user and earn a net reward of 1. An optimal offline algorithm would earn total reward T on *sequence B*; hence, an online algorithm that always matches an arriving costumer if possible can never have a competitive ratio better than $O\left(\frac{1}{T}\right)$.

Let c denote the capacity of the resource. For the general case, consider the following family of arrival sequences and subsequent lemma.

• *Sequence C* (c, t): cT^t users, each with identical usage duration distribution, where the item is either returned immediately with probability $p_t = 1 - \frac{1}{T^t}$ or never returned, that is, $\mathbb{P}(d_t = 0) = p_t = 1 - \frac{1}{T^t}$ and $\mathbb{P}(d_t = \infty) = 1 - p_t = \frac{1}{T^t}$.

In the following, we focus on equitable algorithms that treat all units of the product equally. This simplifies the arguments and is without loss of generality for the overall result. Formally, because all c units of the product are identical, w.l.o.g., each time a unit of the product is to be matched, we let the algorithm randomly pick an available unit for the match. If the algorithm has net expected revenue of at least R, then this ensures that the expected revenue from allocating any individual unit is R/c, since all units are treated equitably. Any algorithm can be turned into an equitable one without change in total revenue.

It is worth noting that an algorithm may for computational reasons differentiate between units of the same resources. Indeed, subsequent work (Goyal et al. 2021) introduces this idea and demonstrates that such differentiation can help in addressing reusability. Our notion of equity applies not at the computational level but to the final allocation. Borrowing an algorithmic idea from Goyal et al. (2021), we explain this in more detail through an example. Consider an instance with two identical units of a single reusable resource that we refer to as *units A* and *B* and a sequence of arrivals all requiring a unit of the resource. Suppose we have an algorithm that computationally maintains a state $(s_1(t), s_2(t)) \in \{0,1\}^2$ and decides whether to allocate a unit of the resource to arrival t based on the state. Initially, $(s_1(1), s_2(1)) = (1,1)$ and we ensure that $s_1(t) + s_2(t) \geq 1$ only if at least one unit of the resource is available at t. The following describes the state-based allocation rule and corresponding state update: (i) if the state at t is (1, 1) or (1, 0), the algorithm allocates a unit at t and updates $s_1(t+1) = 0$; (ii) if the state is (0, 1), then with probability (w.p.) 0.5 it allocates a unit and updates $s_2(t+1) = 0$ and w.p. 0.5 it rejects t; (iii) in state (0, 0), it rejects t. When a unit returns from use at t and the unit was allocated while in state (1, 1) or (1, 0), we update $s_1(t) = 1$. In all other cases, when a unit returns at t, we update $s_2(t) = 1$. Observe that both state (1, 0) and state (0, 1) indicate that exactly one unit is available, but the algorithm behaves differently in the two scenarios. However, this is a purely computational differentiation. Whenever the algorithm makes an allocation, if we have both A and B available, we pick one for allocation uniformly randomly. Thus, the expected number of times A and B are matched is the same, and the algorithm is equitable. With this in mind, consider the following lemma.

Lemma 4. *Given capacity* $c \geq 1$, $t \in [T]$ *and arrival sequence C(c, t), if an equitable algorithm generates expected revenue of at least* $c\frac{1-p_t^{\alpha T^t}}{1-p_t}$ *for some* $\alpha \in [0,1]$, *then for every individual unit of the product, the probability that the unit is consumed forever after the last arrival is at least* $1 - p_t^{\alpha T^t}$.

Proof. Suppose that a single unit, i, is attempted to be matched y times, that is, unit i is matched repeatedly every time it returns from a finite use for up to y times in total. Then the expected total reward from matching i is

$$R_i(y) = (1 - p_t) + 2p_t(1 - p_t) + \cdots + yp_t^{y-1} = \frac{1 - p_t^y}{1 - p_t}.$$

Observe that the expected reward is $\frac{1}{1-p_t}$ times the probability $1-p_t^y$ that the unit is consumed forever (extinguished) when matched up to y times. Now, for any unit i of the resource, the maximum number, y, of match attempts is a random variable. Formally, define random variable $\mathbf{Y}_i$ as the number of times unit i is matched given that the unit always has a finite usage duration. Furthermore, we independently sample usage durations for i and let $\boldsymbol{\tau}_i$ denote the (random) number of finite usage durations before a duration of $+\infty$. Clearly, $\mathbf{Y}_i$ is independent of $\boldsymbol{\tau}_i$, and therefore, the expected revenue from matching unit i in the algorithm is

$$\begin{aligned}\mathbb{E}[R_i(\mathbf{Y}_i,\boldsymbol{\tau}_i)] &= \sum_y \mathbb{P}[\mathbf{Y}_i = y]\,\mathbb{E}[R_i(y,\boldsymbol{\tau}_i)]\\ &= \sum_y \mathbb{P}[\mathbf{Y}_i = y]\frac{\mathbb{P}[i \text{ extinguished}\,|\,\mathbf{Y}_i = y]}{1-p_t}\\ &= \frac{\mathbb{P}[i \text{ extinguished}]}{1-p_t}.\end{aligned}$$

So if the expected revenue from matching unit i in the algorithm is at least $\frac{1-p_t^{\alpha T^t}}{1-p_t}$, then the probability that i is extinguished is at least $1-p_t^{\alpha T^t}$. To complete the proof, recall that in an equitable algorithm, the expected revenue from an individual unit is $1/c$ fraction of the total expected revenue. Given an equitable algorithm with total revenue of at least $c\,\frac{1-p_t^{\alpha T^t}}{1-p_t}$, we have revenue of at least $\frac{1-p_t^{\alpha T^t}}{1-p_t}$ from an individual unit. So the probability that any given unit i survives is at most $p_t^{\alpha T^t}$. Observe that for $T \to \infty$, $1-p_t^{\alpha T^t} \to 1-e^{-\alpha}$. □

Corollary 3. *For any given capacity c and $t \in [T]$, the maximum expected revenue generated by any algorithm (online or offline) on arrival sequence $C(c,t)$ is at least $c\,\frac{1-p_t^{T^t}}{1-p_t}$ and at most cT^t.*

Proof. Clearly, the maximum revenue is at most cT^t. For the lower bound, consider the algorithm that attempts to match each unit of the product to T^t arrivals. From the analysis of Lemma 4, we have that a single unit of the product generates maximum expected revenue $\frac{1-p_t^{T^t}}{1-p_t} = \Theta(T^t)$. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. For arbitrary capacity $c \geq 1$, consider T sequences $D(t) = \{C(c,1),\ldots,C(c,T)\}$ for $t \in [T]$ that begin with cT users arriving from sequence $C(c,1)$, followed by cT^2 users from sequence $C(c,2)$ and so on in order till $C(c,T)$. For any sequence $D(t)$, the maximum possible expected revenue is $\Theta(cT^t)$, because it is lower bounded by $c(1-p_t^{T^t})T^t = \Omega(cT^t)$ (matching only the users in $C(c,t)$ while ignoring earlier users and using Corollary 3) and is upper bounded by $c\left(\sum_{k=1}^t T^t\right) = O(cT^t)$.

We prove by contradiction. Consider a β-competitive online algorithm and assume that $\beta = \Omega(\frac{\log T}{T})$ (otherwise we are done). W.l.o.g., let the algorithm be equitable toward units of the products. From the assumption on competitiveness and using Corollary 3, on arrival sequence $D(1)$ the expected revenue of the online algorithm from an individual unit must be at least $\beta(1-p_1^T)T$. From Lemma 4, we have the probability the unit available after all arrivals is at most $1-\beta(1-p_1^T) \to 1-\beta(1-1/e)$. Now, similar to case of $D(1)$, in order to be β-competitive on sequence $D(2)$, where the maximum expected profit is $\Theta(cT^2)$, the expected reward generated from the $C(c,2)$ part of sequence $D(2)$ must be at least $\beta c(1-p_2^{T^2})T^2$, because the contribution from arrivals $C(c,1)$ is at most $\Theta(cT) = c \times o(\beta T^2)$ for $\beta = \Omega(\frac{\log T}{T})$. Focusing again on an individual unit and applying Lemma 4, the probability of the unit surviving after all arrivals from the $C(c,2)$ part of sequence $D(2)$, conditioned on the unit surviving after arrivals from the $C(c,1)$ part of $D(2)$, is at most $1-\beta(1-p_2^{T^2}) \to 1-\beta(1-1/e)$. Thus, the probability of the unit surviving after all arrivals in $D(2)$ is at most $(1-\beta(1-1/e))^2$. More generally, it follows that the probability of an individual unit surviving after arrivals from sequence $D(t)$ is at most $(1-\beta(1-1/e))^t$. Therefore, on sequence $D(T)$, there is at most a $(1-\beta(1-1/e))^{T-1}$ probability that an individual unit survives until the first arrival from the $C(c,T)$ part of $D(T)$. Hence, the overall expected revenue on $D(T)$ is $c \times O(\max\{(1-\beta(1-1/e))^{T-1}T^T, T^{T-1}\})$. Therefore, the competitive ratio β of the algorithm must satisfy $\beta \leq O(\max\{(1-\beta(1-1/e))^{T-1}, \frac{1}{T}\})$. This translates to $\beta \leq O(\frac{1}{T})$ for $\beta \geq \frac{2\log T}{T}$, a contradiction. Therefore, β is no larger than $\frac{2\log T}{T}$. Note that a more refined argument can be used to further tighten the log factor. □

Remark. Subsequent work (Goyal et al. 2021) shows that no meaningful competitive ratio result is possible even in the case of deterministic user type-dependent usage durations.

4. Computational Experiments

We compare the performance of our myopic policy against the approximate dynamic program-based algorithm (*theDP − basedpolicy*) in Rusmevichientong et al. (2020) and the inventory-balancing policy (*theIBpolicy*) in Golrezaei et al. (2014).

4.1. Experimental Setup

We consider $n = 5$ products indexed by $\mathcal{N} = \{1, \ldots, 5\}$ and $M = 5$ customer types. We consider a selling horizon of $T = 300$ periods. In each period, a random customer from a known distribution over M types arrives. We offer an assortment to each customer when they arrive, who in turn either purchases a product in the offered assortment or leaves the platform without making any purchase. We experiment with three different levels of starting inventory for all products: i) scarce inventory, 1 unit per product; ii) moderate inventory, 5 units per product; and iii) abundant inventory, 20 units per product, to demonstrate the performance of the algorithms at different levels of abundance of products.

We first experiment when the price and usage time distribution for each product is fixed. In the later part of this section, we change the setting parameters to further investigate the performance of the myopic policy when these assumptions do not hold. We select prices for the products to be evenly spaced in $[15, 30]$. Revenue r_i is collected whenever a product of type i is chosen by a customer. The usage time distribution for each product i is a geometric distribution with parameter $p_i \in [0.05, 0.07]$ and expected usage time between 14 and 20 days (until we remove this assumption later). In particular, for product type i, the parameter is $p_i = \frac{1}{20-i}$. Therefore, type 1 product is the most expensive and has the longest expected usage time, and type 5 product is the least expensive and has the shortest expected usage time.

We follow the MNL model with consideration sets in the computational experiments as in Golrezaei et al. (2014) and Rusmevichientong et al. (2020). For any $j \in [M]$, customers of type j have the consideration set $\mathcal{C}_j = \{1, \ldots, j\}$. Each customer makes a choice among the assortment they are provided according to the multinomial logit model. A customer of type j associates the preference weight w_i^j with product i and the preference weight w_0^j with the no-purchase option. When offered assortment S, a customer of type j chooses product $i \in S$ with probability $\phi_i^j(S) = \frac{w_i^j}{w_0^j + \sum_{\ell \in S} w_\ell^j}$. The weight w_i^j over product i of type j customers is generated uniformly randomly from $[0, 1]$ for all j and for any product i in the customer's consideration set. We calibrate the preference weight of the no-purchase option so that for any customer type, if we offer all the products to the customer, the probability of the customer leaving without making a purchase is 0.1, that is, $w_0^j / (w_0^j + \sum_{\ell \in [N]} w_\ell^j) = 0.1$.

We experiment in a scenario that is less favorable to our myopic policy. More specifically, our setting generates customer arrivals so that the pickier customers are more concentrated in the later part of the selling horizon. Therefore, being myopic can be harmful in this setting. We choose equally spaced time periods $\tau^N \le \tau^{N-1} \ldots \le \tau^2 \le \tau^1$ over the selling horizon as in Rusmevichientong et al. (2020). The probability $p^{t,j}$ that a customer of type j arrives at time period t is proportional to $e^{-\kappa|t-\tau^j|}$. So, the arrival probability for a customer of type j peaks at around time period τ^j. Because $\tau^N \le \tau^{N-1} \ldots \le \tau^2 \le \tau^1$, as $\kappa \to \infty$, we obtain an arrival process where customers of type 5 arrive first, followed by customers of type 4. As $\kappa \to 0$, we have $p^{t,j} \to 1/M$, in which case different customer types arrive with equal probability at each time period. Thus, we can control how much the arrival order for the customer types deviates from the equal probability distribution through the parameter κ. For our experiments, we use $\kappa = 0.5$.

4.2. Algorithms and Benchmark

We evaluate the performance of our myopic policy as well as the DP policy and the IB policy. Note that we compute the DP policy assuming the knowledge of the distribution of arrival types, whereas our policy and the IB policy are agnostic to the arrival distribution. We also compare the performance of our myopic policy with the DP-based policy when the realized distribution of arrivals is slightly perturbed. In particular, we consider the following notion of "noise."

We use a scalar λ to control the noise or perturbation from the assumed distribution as follows. At any time t, with probability $1 - \lambda$, the arrival customer type is chosen according to the original distribution (where $p^{t,j} \propto e^{-\kappa|t-\tau^j|}$, $\forall j$), and with probability λ, the arrival customer type is chosen from the uniform probability distribution, where $p^{t,j} = 1/M$. Therefore, when $\lambda = 0$, the $DP-basedpolicy$ has fully accurate distributional knowledge. We evaluate and compare the performances of our myopic algorithm, the DP-based algorithm, and the IB algorithm for different noise levels.

4.3. LP Upper Bound

We use a natural adaptation of the LP upper bound in Dickerson et al. (2018) as a benchmark. This LP imposes inventory constraints in expectation over randomness in rental times:

$$\begin{aligned} \underset{y_{S,t}}{\text{maximize}} \ & \sum_{t=1}^{T} \sum_{S \in \mathcal{S}} \sum_{i=1}^{n} r_i \phi_i^{z_t}(S) y_{S,t} \\ \text{subject to} \ & \sum_{\tau=1}^{t} \sum_{S \in \mathcal{S}} (1 - F_i(t-\tau)) \phi_i^{z_\tau}(S) y_{S,\tau} \\ & \le c_i, i \in [n], t \in [T] \\ & \sum_{S \in \mathcal{S}} y_{S,t} \le 1, t \in [T] \end{aligned} \tag{2}$$

The decision variables $\{y_{S,t}\}$ correspond to the probability that assortment S is offered to the customer arriving at time t.

4.4. Results

For each experiment setting described above or later in this section, we ran the LP benchmark and each algorithm 1,000 times over randomly generated customer preference weights, customer arrivals, usage times, and purchase choices for statistical significance.

Table 1 shows our computational results in the basic setting under different levels of noise (λ) added to the customer arrivals. The columns corresponding to each algorithm give the ratios of the average revenue of the algorithm (more than 1,000 instances) divided by the average revenue of the LP upper bound (more than 1,000 instances) under low-, moderate-, and high-inventory scenarios. The standard deviations of the algorithms are very similar and, therefore, omitted.

Table 1 shows that our myopic algorithm achieves comparable optimality with the other two benchmark algorithms in Rusmevichientong et al. (2020) and in Golrezaei et al. (2014) in practice when the customer arrivals are stochastic and not adversarial, even though our policy is myopic. The myopic policy slightly outperforms the other algorithms when inventory is abundant.

Next, we examine the performance of the myopic policy in more general settings where the analytical guarantee does not hold. Table 2 compares the cumulative revenue of the three algorithms and the LP upper bound when the usage time depends on the users. In particular, every time we run the experiment, for each product-user-type pair, with probability 0.5 the usage time distribution is a geometric distribution as before, and with remaining probability 0.5 this user type never returns this product type. The DP and LP benchmarks incorporate the knowledge of the user-type-dependent usage time distribution in rewards, whereas the myopic policy does not.

Table 2 shows that the DP-based algorithm noticeably outperforms the other two algorithms when the usage time depends on the user type. The ratio between any of the three algorithms and the LP upper bound drops significantly. As we can see, removing the assumption that the usage time for each product is independent of customer type leaves the myopic policy at a disadvantage compared with the DP-based algorithm. However, the difference is not enormous.

Table 3 compares the cumulative revenue of the three algorithms and the LP upper bound when the revenue collected for each product depends on the users. In particular, every time we run the experiment, we generate a uniformly random discount factor for each user type from $[0,1]$, and the revenue we collect from each user is discounted by the user's discount factor.

The performances of the three algorithms stay comparable even with customer-dependent product revenues. Overall, the DP-based algorithm has an advantage over the other two algorithms. However, this advantage diminishes when the inventory level is abundant. We remark that other types of revenue dependence on user types could potentially lead to varied results.

5. Conclusions

In this paper, we consider an online assortment optimization problem with reusable resources or products under an adversarial arrival model. Under the

Table 1. Comparison Under Different Levels of Noise (Higher λ Implies More Noise) Added to the Customer Arrivals, with Inventory Levels C = 1, 5, 20, and Arrival Distribution Parameter $\kappa = 0.5$

	Myopic Alg			*IB Alg*			*DP-based Alg*		
λ	$C=1$	$C=5$	$C=20$	$C=1$	$C=5$	$C=20$	$C=1$	$C=5$	$C=20$
0.0	0.949	0.927	0.996	0.950	0.927	0.994	0.948	0.937	0.975
0.2	0.908	0.913	0.995	0.909	0.911	0.996	0.913	0.931	0.973
0.5	0.909	0.895	0.994	0.908	0.893	0.992	0.926	0.917	0.970
1	0.910	0.855	0.993	0.914	0.857	0.994	0.919	0.884	0.970

Notes. An entry represents the average performance of the algorithm divided by the average value of the LP upper bound. The standard deviations of revenue for the three algorithms in the 1,000 repeated experiments for each inventory level are very similar and, therefore, omitted.

Table 2. Comparison When Usage Time Depends on the Customer Type, with C = 1, 5, 20

	Myopic Alg			*IB Alg*			*DP-based Alg*		
λ	$C=1$	$C=5$	$C=20$	$C=1$	$C=5$	$C=20$	$C=1$	$C=5$	$C=20$
0.0	0.522	0.480	0.666	0.514	0.479	0.664	0.521	0.487	0.709
0.2	0.376	0.436	0.694	0.377	0.435	0.694	0.382	0.443	0.737
0.5	0.318	0.405	0.707	0.313	0.404	0.708	0.328	0.413	0.743
1	0.315	0.397	0.750	0.314	0.396	0.750	0.317	0.408	0.780

Note. The standard deviations of revenue for the three algorithms in the 1,000 repeated experiments are very similar.

Table 3. Comparison When Revenue Depends on the Customer Type, with C = 1, 5, 20

	Myopic Alg			*IB Alg*			*DP-based Alg*		
λ	$C=1$	$C=5$	$C=20$	$C=1$	$C=5$	$C=20$	$C=1$	$C=5$	$C=20$
0.0	0.936	0.924	1.000	0.941	0.924	1.000	0.929	0.931	0.982
0.2	0.794	0.893	0.989	0.790	0.896	0.991	0.801	0.909	0.970
0.5	0.769	0.868	0.974	0.774	0.867	0.972	0.780	0.892	0.951
1	0.752	0.840	0.970	0.748	0.840	0.969	0.770	0.868	0.947

Note. The standard deviations of revenue for the three algorithms in the 1,000 repeated experiments are very similar.

assumption that product usage time distributions do not depend on the user type, we show that the policy that offers a myopically optimal assortment to every user from the set of available products achieves an expected revenue that is at least $1/2$ times the expected revenue of a clairvoyant algorithm that has full information about the sequence of user types. For the case of reusable capacities, we do not have a good upper bound (LP based or otherwise) for the clairvoyant optimal, which makes the comparison with the benchmark challenging. The main contribution of this paper is a queue-based coupling technique that allows us to relate the expected revenue of the clairvoyant optimal to the expected revenue of the myopic policy. This coupling is algorithmic and might be of independent interest. The assumption that product usage time distribution does not depend on user type is fairly reasonable and satisfied in many settings. We also show that if the assumption is not satisfied, there is no online algorithm that can be constant-factor competitive as compared with our clairvoyant benchmark. Therefore, the assumption is necessary to get any nontrivial performance guarantee for the case of adversarial arrivals.

Our myopic online algorithm is easy to implement and achieves comparable optimality with the DP-based algorithm in Rusmevichientong et al. (2020) and the inventory-balancing algorithm in Golrezaei et al. (2014) in synthetic experiments, even when some of the assumptions we make in this paper do not hold.

An interesting open question is to study whether we can obtain stronger results analogous to the online assortment problem with non-reusable and large capacities, in particular, a $(1-1/e)$-competitive algorithm in the adversarial arrivals model and a near-optimal result in the stochastic arrivals model. The first of these open questions was subsequently resolved in Goyal et al. (2021) (which merges Goyal et al. (2020a) and Goyal et al. (2020b)) and the second in Feng et al. (2020).

References

Baek J, Ma W (2019) Bifurcating constraints to improve approximation ratios for network revenue management with reusable resources. Preprint, submitted December 15, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3482457.

Ball MO, Queyranne M (2009) Toward robust revenue management: Competitive analysis of online booking. *Oper. Res.* 57(4): 950–963.

Berbeglia G, Garassino A, Vulcano G (2021) A comparative empirical study of discrete choice models in retail operations. *Management Sci.* Forthcoming.

Bernstein F, Kök AG, Xie L (2015) Dynamic assortment customization with limited inventories. *Manufacturing Service Oper. Management* 17(4):538–553.

Blanchet JH, Gallego G, Goyal V (2016) A Markov chain approximation to choice modeling. *Oper. Res.* 64(4):886–905.

Budhiraja A, Mukherjee D, Wu R (2019) Supermarket model on graphs. *Ann. Appl. Probab.* 29(3):1740–1777.

Chan CW, Farias VF (2009) Stochastic depletion problems: Effective myopic policies for a class of dynamic optimization problems. *Math. Oper. Res.* 34(2):333–350.

Chen Y, Levi R, Shi C (2017) Revenue management of reusable resources with advanced reservations. *Production Oper. Management* 26(5):836–859.

Cruise J, Jonckheere M, Shneer S (2020) Stability of jsq in queues with general server-job class compatibilities. *Queueing Systems* 95:271–279.

Davis JM, Gallego G, Topaloglu H (2014) Assortment optimization under variants of the nested logit model. *Oper. Res.* 62(2): 250–273.

Devanur NR, Jain K, Kleinberg RD (2013) Randomized primal-dual analysis of ranking for online bipartite matching. *Proc. 24th Annual ACM-SIAM Sympos. Discrete Algorithms SODA '13* (USA: Society for Industrial and Applied Mathematics), 101–107.

Dickerson JP, Sankararaman KA, Srinivasan A, Xu P (2018) Allocation problems in ride-sharing platforms: Online matching with offline reusable resources. McIlraith SA, Weinberger KQ, eds. *Proc. 32nd AAAI Conf. Artificial Intelligence (AAAI-18)* (AAAI Press), 1007–1014.

Feng Y, Niazadeh R, Saberi A (2019) Linear programming based online policies for real-time assortment of reusable resources. Preprint, submitted July 17, https://dx.doi.org/10.2139/ssrn.3421227.

Feng Y, Niazadeh R, Saberi A (2020) Near-optimal bayesian online assortment of reusable resources. Preprint, submitted October 21, https://ssrn.com/abstract=3714338.

Gallego G, Iyengar G, Phillips R, Dubey A (2004) Managing flexible products on a network. Preprint, submitted April 27, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3567371.

Gallego G, Topaloglu H (2014) Constrained assortment optimization for the nested logit model. *Management Sci.* 60(10):2583–2601.

Golrezaei N, Nazerzadeh H, Rusmevichientong P (2014) Real-time optimization of personalized assortments. *Management Sci.* 60 (6):1532–1551.

Goyal V, Iyengar G, Udwani R (2020a) Online allocation of reusable resources: Achieving optimal competitive ratio. Preprint, submitted February 6, https://arxiv.org/abs/2002.02430.

Goyal V, Iyengar G, Udwani R (2020b) Online allocation of reusable resources via algorithms guided by fluid approximations. Preprint, submitted October 8, https://arxiv.org/abs/2010.03983.

Goyal V, Iyengar G, Udwani R (2021) Asymptotically optimal competitive ratio for online allocation of reusable resources. Preprint, submitted February 6, https://arxiv.org/abs/2002.02430.

Goyal V, Udwani R (2020) Online matching with stochastic rewards: Optimal competitive ratio via path based formulation. *Proc. 21st ACM Conf. Econom. Comput. EC '20* (Association for Computing Machinery, New York), 791.

Karp RM, Vazirani UV, Vazirani VV (1990) An optimal algorithm for on-line bipartite matching. *Proc. 22nd Annual ACM Sympos. Theory Comput., STOC '90* (Association for Computing Machinery, New York), 352–358.

Kök A, Fisher M, Vaidyanathan R (2015) Assortment planning: Review of literature and industry practice. Agrawal N, Smith SA, eds. *Retail Supply Chain Management, volume 223 of International Series in Operations Research & Management Science* (Springer, New York), 175–236.

Levi R, Radovanović A (2010) Provably near-optimal lp-based policies for revenue management in systems with reusable resources. *Oper. Res.* 58(2):503–507.

Liu Q, van Ryzin GJ (2008) On the choice-based linear programming model for network revenue management. *Manufacturing Service Oper. Management* 10(2):288–310.

Luce RD (1959) *Individual Choice Behavior: A Theoretical Analysis* (Wiley, New York).

Ma W, Simchi-Levi D (2020) Algorithms for online matching, assortment, and pricing with tight weight-dependent competitive ratios. *Oper. Res.* 68(6):1787–1803.

McFadden D (1973) Conditional logit analysis of qualitative choice behavior. Zarembka P, ed. *Frontiers in Econometrics* (Academic Press, New York), 105–142.

McFadden D (1978) Modelling the choice of residential location. Karlqvist A, Lundqvist L, Snickars F, Weibull J, eds. *Spatial Interaction Theory and Planning Models* (North Holland, Amsterdam), 75–6.

McFadden D, Train K (2000) Mixed MNL models for discrete response. *J. Appl. Econometrics* 15(5):447–470.

Mehta A (2013) Online matching and ad allocation. *Foundation Trends Theoret. Comput. Sci.* 8(4):265–368.

Mukherjee D, Borst SC, van Leeuwaarden JSH (2018) Asymptotically optimal load balancing topologies. *Proc. ACM Measurement Anal. Comput. Systems* 14:1–29.

Owen Z, Simchi-Levi D (2018) Price and assortment optimization for reusable resources. Preprint, revised February 20, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3070625.

Plackett RL (1975) The analysis of permutations. *J. Roy. Statist. Soc.: Ser. C. Appl. Statist.* 24(2):193–202.

Rusmevichientong P, Sumida M, Topaloglu H (2020) Dynamic assortment optimization for reusable products with random usage durations. *Management Sci.* 66(7):2820–2844.

Stein C, Truong VA, Wang X (2018) Advance service reservations with heterogeneous customers. *Management Sci.* 66(7):2801–3294.

Talluri K, Van Ryzin G (2004) Revenue management under a general discrete choice model of consumer behavior. *Management Sci.* 50(1):15–33.

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

Train KE (2009) *Discrete Choice Methods with Simulation* (Cambridge University Press, Cambridge, UK).

Wang X, Truong VA, Bank D (2018) Online advance admission scheduling for services with customer preferences. Preprint, submitted May 26, https://arxiv.org/abs/1805.10412.

Weng W, Zhou X, Srikant R (2020) Optimal load balancing with locality constraints. *Proc. ACM Measurement Anal. Comput. Systems* 45:1–37.

Williams HW (1977) On the formation of travel demand models and economic evaluation measures of user benefit. *Environ. Planning A* 3(9):285–344.