

On Policies for Single-leg Revenue Management with Limited Demand Information

Will Ma

Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139, willma@mit.edu

David Simchi-Levi

Institute for Data, Systems, and Society, Department of Civil and Environmental Engineering, and Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139, dslevi@mit.edu

Chung-Piaw Teo

NUS Business School, National University of Singapore, Singapore, bizteocp@nus.edu.sg

In this paper we study the single-leg revenue management problem, with no information given about the demand trajectory over time. The competitive ratio for this problem has been established by Ball and Queyranne (2009) under the assumption of independent demand, i.e., demand for higher fare classes does not spill over to lower fare classes. We extend their results to general demand models to account for the buying-down phenomenon, by incorporating the price-skimming technique from Eren and Maglaras (2010). That is, we derive state-dependent price-skimming policies, which stochastically increase their price distributions as the inventory level decreases, in a way that yields the best-possible competitive ratio. Furthermore, our policies have the benefit that they can be easily adapted to exploit available demand information, such as the personal characteristics of an incoming online customer, while maintaining the competitive ratio guarantee. A key technical ingredient in our paper is a new “valuation tracking” subroutine, which tracks the possible values for the optimum, and follows the most inventory-conservative control which maintains the desired competitive ratio.

Key words: online algorithms, competitive ratio, revenue management, dynamic pricing

1. Introduction

In this paper we consider the single-leg revenue management problem, where a firm is selling multiple products that share a single capacity over a finite time horizon. The price of each product

and the unrepensible starting capacity are exogenously determined. The firm's objective is to maximize its total revenue earned, by dynamically controlling the availability of different products over time. The tradeoff lies between “myopic” controls which maximize the immediate revenue at a point in time, and “conservative” controls which ration capacity for the remaining time horizon.

The motivating application for this problem lies in airlines, where each flight leg has a limited seat capacity, and the products correspond to different “fare classes” (e.g. economy, business) which offer seats at different prices. The seat capacity and fare classes have been determined long in advance, through factors such as business strategy, positioning, competition, etc. The time horizon is finite, ending upon the flight's departure.

We study this problem in the setting where very limited information is given or can be learned about demand. This setting was introduced by Ball and Queyranne (2009), Lan et al. (2008), where meaningful controls can be derived based on only the knowledge of the fare class prices. The authors consider *booking limit* policies, which can be described as follows. Initially, all of the fare classes are made available to customers. Once the total seat sales surpasses a critical threshold, the lowest fare becomes unavailable. Progressively higher fares are made unavailable until the flight either becomes full or takes off.

An important assumption made in the model of Ball and Queyranne (2009) is that demand for the different fare classes is *independent*. That is, although the lower fares are made available until their booking thresholds are reached, there is no risk of *cannibalizing* the sales of higher fares. The justification for this assumption in the literature is two-fold. First, the fare classes have been designed to segment customers and achieve price discrimination, i.e. the perks provided by business class dissuade price-insensitive business travelers from having interest in lower-class fares. Second, these business travelers tend to book last-minute, i.e. by the time they book the lower-class fares have usually become unavailable anyway.

However, marketplace and technology changes have introduced environments where these justifications may no longer hold. For example, the competition from low-cost carriers (e.g. Spirit

Airlines) has forced major airlines (e.g. United Airlines) into adopting more overlapping fare family structures (e.g. adding a “basic economy” fare which competes with the regular economy fare). Also, the advent of e-commerce has brought a significant increase in customer sophistication, as well as the firm’s understanding of customers via “big data” analytics. The motivation of this paper is to derive advanced booking controls which rely on dynamic price-changing infrastructure instead of the independence assumption to differentiate between customers with different willingness-to-pay.

1.1. Model and Results

Throughout this paper, we analyze the “critical” case where customers always substitute to the lowest available fare. In Section 6, we explain why this is sufficient to allow for any setting where customers substitute according to a random-utility choice model.

In this critical case, the problem reduces to a *dynamic pricing* problem. We can describe each customer using a *valuation*, or maximum willingness-to-pay. At each point in time a single price is offered, and customers who encounter that price make a purchase if and only if it does not exceed their valuation. (Offering multiple prices is redundant, because the lower price would always be chosen over the higher price.)

We let \mathcal{P} denote the set of fare class prices, which are the feasible prices to charge. The starting capacity comes in the form of k discrete units of *inventory*. The selling horizon consists of T discrete time steps, which are sufficiently granular such that at most one customer arrives during each time step. We let V_t denote the maximum price in \mathcal{P} that the customer in time t is willing to pay, which is 0 if no customer arrived. An *online algorithm* must sequentially choose a price P_t for each time t , and if $V_t \geq P_t$, then revenue P_t is earned and one unit of inventory is depleted.

In settings where no information is given about the sequence of valuations, an online algorithm is evaluated by comparing its total revenue earned on different sequences to that of a clairvoyant optimum. For any sequence V_1, \dots, V_T , the *offline optimum* $\text{OPT}(V_1, \dots, V_T)$ is defined as the maximum revenue that could have been earned from knowing all the valuations in advance, equal to the k largest values in V_1, \dots, V_T . For $c \leq 1$, if an online algorithm can guarantee that its revenue

is at least $c \cdot \text{OPT}(V_1, \dots, V_T)$ on every sequence V_1, \dots, V_T , then it is said to be *c-competitive*. If c is best-possible in that any (potentially randomized) online algorithm cannot simultaneously guarantee greater than $c \cdot \text{OPT}(V_1, \dots, V_T)$ revenue over all sequences V_1, \dots, V_T , then c is called the *competitive ratio*.

In the model of Ball and Queyranne (2009), the competitive ratio for any problem instance is shown to be a function of only the price set \mathcal{P} , which we will denote using $c^*(\mathcal{P})$. Their model corresponds to ours when each V_t is *deterministic* and given at the *start* of time t . The decision at time t then reduces to an accept-reject decision, where accepting customer t corresponds to charging her maximum willingness-to-pay of V_t , and rejecting customer t corresponds to charging any price above V_t (possibly ∞ if the inventory has run out). Ball and Queyranne (2009) derive $c^*(\mathcal{P})$ -competitive booking limit policies which specify when to reject customers paying low prices.

In this paper, we derive $c^*(\mathcal{P})$ -competitive online algorithms under the following models with progressively less information:

1. V_t is *deterministic* and given at the *end* of time t (Section 2). This model includes the case where the airline has a point estimate of V_t at the start of time t , but does not use it in determining the price P_t , because the airline wants to refrain from “personalized pricing”.
2. V_t is *stochastic* and distributionally-given at the *end* of time t (Section 3). In this model, the airline estimates a conditional distribution of the customer’s valuation V_t falling into the different fare classes, by using the arrival patterns of different fare classes as well as the customer’s response to being offered price P_t .
3. No information on V_t is ever given (Section 4).

All of these algorithms are best-possible. Indeed, since an online algorithm cannot be better than $c^*(\mathcal{P})$ -competitive under the model of Ball and Queyranne (2009), an online algorithm also cannot be better than $c^*(\mathcal{P})$ -competitive under our models with less information.

Our algorithms use the *price-skimming* technique of Eren and Maglaras (2010), who analyze how the price of an item should be distributed (e.g. across stores, across time) when the price set \mathcal{P} is

known but demand is completely unknown. Their model corresponds to ours when the inventory constraint is irrelevant (e.g. when $k \geq T$), in which case they show that the competitive ratio is also $c^*(\mathcal{P})$. Our work shows how the price-skimming distribution should depend on inventory when it is relevant. More specifically, our algorithm determines the optimal path-dependent pricing distribution for each time step t , conditional on the amount of remaining inventory. Interestingly, our algorithm has the following feature: at any time step, the price distribution which maximizes the competitive ratio is strictly stochastically-decreasing in the amount of remaining inventory (see Section 2.4). In fact, this is analogous to a classical structural property when the demand sequence is known or distributionally-known: at any time step, the price which maximizes the expected revenue is strictly decreasing in the amount of remaining inventory (Gallego and Van Ryzin 1994, Zhao and Zheng 2000).

Finally, in Section 5, we consider the personalized online revenue management setting where the distribution of each V_t becomes available at the start of time t , and this information can (and should) be used in determining the offering to customer t . Introduced by Golrezaei et al. (2014), this setting has a very practical motivation of e-commerce, where purchase probabilities are estimated for each customer based on her personal characteristics (e.g. Mac users have a higher chance of buying business-class flights). Although our algorithms from Sections 2–3 focus on the setting without personalized pricing, we show in Section 5 that our algorithms can be naturally adapted to exploit personal information. The adapted algorithms maintain the competitive ratio of $c^*(\mathcal{P})$, which is best-possible even with personalized information, since $c^*(\mathcal{P})$ is best-possible under the setting of Ball and Queyranne (2009) where each customer’s personal valuation is known deterministically.

1.2. Sketch of Techniques, and Comparison with Existing Techniques

The main technical contribution behind our results is a new “valuation tracking” procedure which incorporates both booking limits and price-skimming. We motivate it using the following example, under the model where each valuation V_t is deterministic and revealed at the end of time t . The

price set is $\mathcal{P} = \{1, 2, 4\}$, and we will refer to customers with these valuations as being of type-L (Low), type-M (Medium), and type-H (High), respectively.

The competitive ratio for this price set derived by Ball and Queyranne (2009) and Eren and Maglaras (2010) is $c^*(\mathcal{P}) = 1/2$, and we describe below their respective $1/2$ -competitive policies.

- Booking Limits (Ball and Queyranne 2009): Initially charge \$1; increase the price to \$2 after $1/2$ of the starting inventory has been sold; further increase the price to \$4 after $3/4$ of the starting inventory has been sold. (This is the variant of booking limits with “theft nesting”.)
- Price-skimming (Eren and Maglaras 2010): Charge \$1 for $1/2$ of the time steps; charge \$2 for $1/4$ of the time steps; charge \$4 for $1/4$ of the time steps.

We now discuss what happens if we try to implement these policies under our model.

Attempt 1: Direct implementation of booking limits. It is easy to see that this would not be $1/2$ -competitive—suppose just one type-H customer arrived at the start. The algorithm would be charging the low price of \$1, while the offline optimum would be the customer’s valuation of \$4.

Any direct implementation of price-skimming would suffer similarly, since there could be a single type-H customer who arrives during a time when the price is set to \$1.

Attempt 2: Price-skimming as a randomized price. It appears that the problem with Attempt 1 can be solved using the “random price” interpretation of price-skimming—instead of deterministically partitioning the time horizon according to ratios $\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$ and offering prices 1, 2, 4 respectively, one could at each time step choose the prices randomly with respective probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$. Then, if a single type-H customer arrives, the expected revenue would be

$$\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 4 = 2$$

which is $1/2$ of the customer’s valuation of \$4. It can be checked that $1/2$ of the customer’s valuation is also earned when it is \$1 or \$2; this is by construction of the price-skimming distribution.

However, having a fixed price-skimming distribution is no longer effective under inventory constraints. Indeed, if a long sequence of type-L customers arrive, then this would deplete the inventory

with high probability, and type-H customers who arrive last-minute would not be served, and the ratio of optimum earned would again be $1/4$.

Attempt 3: Naive incorporation of booking limits into price-skimming. It appears that the problem with Attempt 2 can be solved by respecting the booking limits, i.e. forbidding price-skimming from randomly choosing the price of \$1 after $1/2$ of the starting inventory has been sold. However, this still fails to be $1/2$ -competitive, as shown by the following example. Suppose the starting inventory is 4, and that 2 type-H customers arrive followed by a type-L customer, with no customers arriving after that. The optimum would be \$9. However, the algorithm's revenue would only equal \$4: it would earn \$2 in expectation from each of the type-H customers, depleting 2 units of inventory, and then earn \$0 from the type-L customer due to the booking limit.

Our procedure: Valuation tracking. The problem with Attempt 3 leads to the following observation—the optimum is guaranteed to increase from the first 4 customers (since there are 4 units of inventory), so in order to be $1/2$ -competitive, the algorithm must maintain the initial price-skimming distribution for the first 4 customers. After that, the algorithm can respect booking limits as long as customers rejected in this way would not increase the optimum, and in fact should do so, to avoid the problem in Attempt 2 of stocking out. This motivates our procedure below.

- Valuation Tracking: At each time t , let ℓ_t denote the smallest value (possibly 0) in the 4 largest valuations to have arrived before time t (since 4 is the starting inventory). Then, randomly choose the price in a way such that the algorithm's expected revenue during time t is *equal to*

$$\frac{1}{2}(V_t - \ell_t), \quad \text{for any } V_t > \ell_t. \quad (1)$$

$V_t - \ell_t$ is the gain in the offline optimum should the valuation of customer t be V_t , and $1/2$ is the desired competitive ratio. The constraint that the algorithm's revenue is exactly equal to (1) effectively forces it to use the most *inventory-conservative* controls which maintains $1/2$ -competitiveness, thereby hedging against a stockout. The price distribution used at each time t depends on the inventory state, and in fact, the calculation for the algorithm's expected revenue

must account for the probability of stocking out before time t . The surprising fact is that it is *possible* to choose price distributions which collectively guarantee (1).

We illustrate our basic valuation tracking procedure in Section 2.1 by “stacking” the past valuations in a geometric configuration, and formalize it in Section 2.2. The basic procedure is designed to have a clean analysis and we consider variants in Sections 2.3–2.4, as well as generalize it to the stochastic-valuation model in Section 3.2. A unique challenge arises in Section 3.3, where we use sampling to convert a $c^*(\mathcal{P})$ -competitive valuation tracking procedure with exponential runtime into a $(c^*(\mathcal{P}) - \varepsilon)$ -competitive procedure with polynomial runtime.

1.3. Related Work

Single-leg revenue management is a cornerstone problem in the area of revenue management and pricing, as outlined in the book by Talluri and Van Ryzin (2006). Many different approaches for modeling demand have been considered over the years, as surveyed in Araman and Caldentey (2011), den Boer (2015). In contrast to other approaches, which optimize based on a forecast of future demand (Gallego and Van Ryzin 1994, Zhao and Zheng 2000), learn an unknown demand function (Besbes and Zeevi 2009), or learn in various Bayesian (Araman and Caldentey 2009, Chen and Wu 2016) and minimax (Lim and Shanthikumar 2007, Zhang et al. 2016) settings, competitive ratio analysis is a form of minimax analysis that operates without demand information.

Our work extends the notion of booking limits from Ball and Queyranne (2009) to allow for non-independent demand, by incorporating the price-skimming of Eren and Maglaras (2010). We should point out that our “random price” interpretation of price-skimming originated from Bergemann and Schlag (2008). Randomization is a powerful technique for improving the competitive ratio of online algorithms; for further background we refer to the book by Borodin and El-Yaniv (2005).

Motivated by e-commerce, the competitive ratio has also been recently studied in many “personalized recommendation” settings (Golrezaei et al. 2014, Chen et al. 2016, Ma and Simchi-Levi

2017, Cheung et al. 2018). In these papers, there are multiple commodities (“legs”) and interchangeability in which commodities are offered to each customer. Our paper differs from these papers in two ways. First, we do not rely on personalized controls, and instead derive a global price-skimming distribution, which is useful for online retailers who choose not to engage in personalized pricing. Also, even when personalization is desired, we show in Section 5 how to effectively use personalized information in the single-leg case, and obtain a competitive ratio $c^*(\mathcal{P})$ which is greater than the multi-leg competitive ratio from Ma and Simchi-Levi (2017).

Finally, while we analyze the problem of inventory-constrained dynamic pricing throughout this paper, our results easily generalize to the dynamic assortment setting, as explained in the conclusion (Section 6) of this paper. Our techniques prescribe the best-possible trade off between revenue maximization and inventory consumption (see Maglaras and Meissner (2006)) when there is a single flight leg, known fare classes, but general unknown demand.

2. Deterministic Valuations

In this section we consider the problem defined in Section 1.1 under the model where each customer’s valuation is deterministic and revealed immediately after she leaves.

First we define some additional notation to that defined in Section 1.1. For any positive integer n , let $[n]$ denote the set $\{1, \dots, n\}$. For notational convenience, we will assume that \mathcal{P} consists of m discrete prices, i.e. $\mathcal{P} = \{r^{(j)} : j \in [m]\}$, sorted $0 < r^{(1)} < \dots < r^{(m)}$. All of our results can be generalized to the case where \mathcal{P} is a continuum of prices taking the form $[r^{\min}, r^{\max}]$, as we discuss in Section D. We define $r^{(0)}$ to be 0, and then the valuation V_t at any time t lies in $r^{(0)}, \dots, r^{(m)}$, with $V_t = r^{(0)}$ representing the lack of a customer during time t . Similarly, we define $r^{(m+1)}$ to be ∞ , and then the price P_t at any time t lies in $r^{(1)}, \dots, r^{(m+1)}$, with $P_t = r^{(\infty)}$ representing the firm shutting off demand during time t , which is the only option if its inventory is out of stock. Let X_t be the indicator variable for making a sale during time t , i.e. it is 1 if $V_t \geq P_t$, and 0 otherwise.

Let T denote the number of time steps. None of the algorithms in this paper assume any knowledge of T ; note that T can always be made arbitrarily large by inserting customers with valuation 0.

We will hereafter treat T as the unknown total number of customers, and use the phrase “customer t ” to refer to valuation V_t (even if it is 0).

An *online algorithm* must choose each P_t based on the history of past prices and valuations, $P_1, V_1, \dots, P_{t-1}, V_{t-1}$. This history also determines the values of X_1, \dots, X_{t-1} . The online algorithm does not know T , and has no information about V_t, V_{t+1}, \dots, V_T , when choosing P_t . In contrast, the *offline optimum* knows the entire sequence V_1, \dots, V_T before having to choose any prices. Given any valuation sequence V_1, \dots, V_T , we use the P_t and X_t variables to refer to the execution of an online algorithm on the valuation sequence. Since the online algorithm may be randomized, we treat P_t and X_t as random variables. Let $\text{ALG}(V_1, \dots, V_T)$ denote the total revenue earned by the online algorithm, equal to $\sum_{t=1}^T P_t X_t$. Then $\mathbb{E}[\text{ALG}(V_1, \dots, V_T)]$ is its expected revenue. Meanwhile, let $\text{OPT}(V_1, \dots, V_T)$ denote the offline optimum for sequence V_1, \dots, V_T , equal to the $\min\{k, T\}$ largest valuations from V_1, \dots, V_T . Formally, an online algorithm is said to be c -competitive if

$$\mathbb{E}[\text{ALG}(V_1, \dots, V_T)] \geq c \cdot \text{OPT}(V_1, \dots, V_T), \quad \forall T \geq 1, (V_1, \dots, V_T) \in (\mathcal{P} \cup \{0\})^T. \quad (2)$$

We will omit the arguments (V_1, \dots, V_T) in ALG and OPT when the context is clear.

As explained in Section 1.1, since our problem captures the problems of both Ball and Queyranne (2009) and Eren and Maglaras (2010), an upper bound for the value of c in (2) is given by $c^*(\mathcal{P})$, as defined below.

DEFINITION 1. For any $m \geq 1$, $0 < r^{(1)} < \dots < r^{(m)}$, and $\mathcal{P} = \{r^{(1)}, \dots, r^{(m)}\}$, define:

- $q^{(j)} = 1 - \frac{r^{(j-1)}}{r^{(j)}}$ for all $j \in [m]$ (recall that $r^{(0)} = 0$);
- $q = \sum_{j=1}^m q^{(j)}$;
- $c^*(\mathcal{P}) = \frac{1}{q}$.

The interpretation of $q^{(j)}$ in Ball and Queyranne (2009) is the fraction of initial inventory “set aside” for prices j and higher. The interpretation of $q^{(j)}$ in Eren and Maglaras (2010) is the fraction of time that price j should be charged. Both of the papers establish that the competitive ratio cannot be better than $c^*(\mathcal{P})$ via Yao’s minimax principle (Yao 1977). Therefore, for any fixed

(but possibly randomized) online algorithm in our problem, there exists a sequence V_1, \dots, V_T such $\mathbb{E}[\text{ALG}(V_1, \dots, V_T)] \leq c^*(\mathcal{P}) \cdot \text{OPT}(V_1, \dots, V_T)$. In this paper we derive various $c^*(\mathcal{P})$ -competitive algorithms, using our valuation tracking procedure as the core subroutine.

2.1. Intuition behind Valuation Tracking Procedure

The goal of our basic procedure is to, for each customer, earn a constant fraction $c^*(\mathcal{P})$ of the gain in OPT from that customer arriving, which would imply being $c^*(\mathcal{P})$ -competitive. To accomplish this, it tracks the current value of OPT , i.e. the sum of the k largest valuations to have arrived thus far, which are assumed to be known.

Consider an example where the feasible price set is $\mathcal{P} = \{1, 2, 4\}$, in which case $c^*(\mathcal{P}) = \frac{1}{2}$. Suppose the starting inventory is $k = 5$, and that 5 customers, with valuations 4, 1, 4, 1, 2, have already arrived. The current value of OPT is then the sum of these 5 valuations, $4 + 1 + 4 + 1 + 2 = 12$.

The procedure considers the possibilities for the increase in OPT from the next customer, which we denote as ΔOPT . Since the smallest valuation currently counted toward OPT is 1, if the valuation of the next customer is 4, then $\Delta\text{OPT} = 3$; if it is 2, then $\Delta\text{OPT} = 1$. If the next customer has valuation not exceeding 1, then $\Delta\text{OPT} = 0$. The procedure wants to guarantee that its *expected* revenue on the next customer is at least $\frac{1}{2} \cdot \Delta\text{OPT}$, for all of these possible valuations. To accomplish this, it has to consider the probability that it has stocked out at this point; on those sample paths its revenue is 0.

Our procedure cleanly accounts for the probability of stocking out using the following approach. Each customer is assigned to a *specific* unit of inventory $i \in [k]$ upon arrival. Each inventory unit i maintains a variable `level`[i], which is the maximum valuation of a customer previously assigned to it. The next customer is always assigned to an unit i^* with the smallest value of `level`[i^*], *regardless of whether that unit i^* has already been sold*. In this way, the assignment procedure is deterministic, and allows us to maintain an invariant: the probability a unit i has been sold is dependent on only the (deterministic) value of `level`[i].

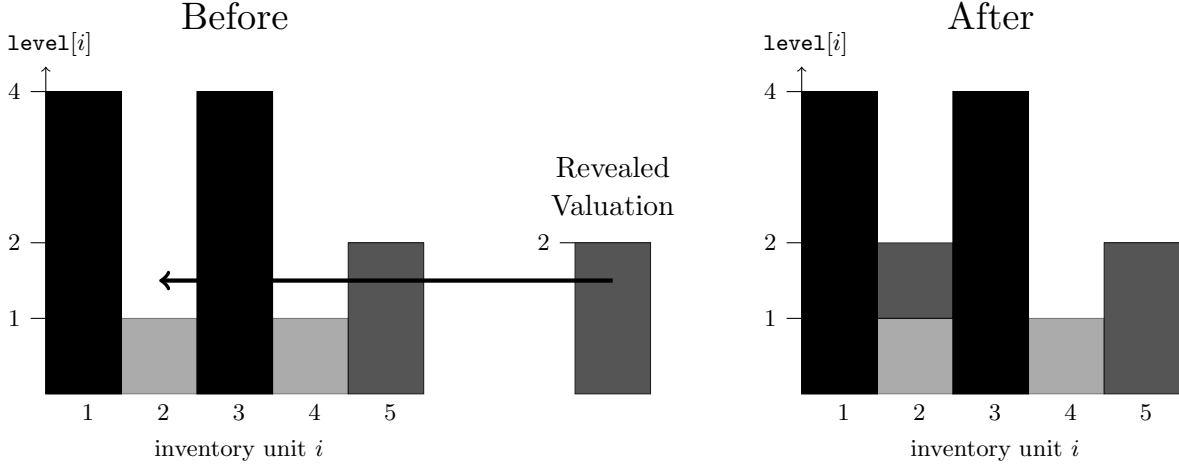


Figure 1 The configuration of valuations, before and after a customer with valuation 2 arrives.

For each customer, the procedure makes an offer to her *only if unit i^* has not been sold*, at a random price exceeding $\text{level}[i^*]$. The higher $\text{level}[i^*]$ is, the more likely it is that unit i^* has been sold, and the lower the expected revenue from that customer. However, if $\text{level}[i^*]$ is high, then the potential increase in OPT from that customer is also lower; if the valuation of the customer does not exceed $\text{level}[i^*]$, then both the procedure's revenue and ΔOPT are 0. By properly choosing the distributions for the random prices, our procedure is able to maintain the invariant on the probability of each unit being sold, while earning $\frac{1}{2} \cdot \Delta\text{OPT}$ in expectation from each customer.

Returning to the example, given that the first 5 customers had valuations 4, 1, 4, 1, 2, the values of $\text{level}[i]$ for $i = 1, \dots, k$ are shown in the LHS of Fig. 1. The next customer, “customer #6”, is assigned to inventory unit 2. After her valuation is revealed to be 2, the updated configuration is shown on the RHS of Fig. 1, regardless of whether she was rejected.

Customer #6 would have been rejected if unit 2 was sold before her arrival, even if other units were available. When $\mathcal{P} = \{1, 2, 4\}$, the probability that a unit i has been sold equals $0, \frac{1}{2}, \frac{3}{4}, 1$ if $\text{level}[i]$ is 0, 1, 2, 4, respectively. These probabilities correspond to the values of $q^{(j)}$ from Definition 1. Since $\text{level}[2]$ was 1 before customer #6 arrived, she is made an offer with probability $\frac{1}{2}$, at a random price exceeding 1. The price is 2 with probability proportional to $\frac{1}{4}$, and 4 with probability proportional to $\frac{1}{4}$ (again using the values of $q^{(j)}$), hence each price would be offered

Algorithm 1: Valuation Tracking Procedure

Input: Customers $t = 1, 2, \dots$ arriving online, with each valuation V_t revealed after the price P_t is chosen.

Output: For each customer t , a (possibly random) price P_t for her.

```

1 level[ $i$ ] = 0, sold[ $i$ ] = false for  $i = 1, \dots, k$ ;
2  $t = 1$ ;
3 while customer  $t$  arrives do
4    $i_t^* = \arg \min_i \text{level}[i]$ ;
5   set  $\ell_t$  to the index in  $\{0, \dots, m\}$  such that  $\text{level}[i_t^*] = r^{(\ell_t)}$ ;
6   if sold[ $i_t^*$ ] = false then
7     offer price  $r^{(j)}$  with probability  $\frac{q^{(j)}}{\sum_{j'=\ell_t+1}^m q^{(j')}}$ , for all  $j = \ell_t + 1, \dots, m$ ;
8   else
9     reject the customer by choosing price  $\infty$ ;
10  end
11  observe valuation  $V_t$  and purchase decision  $X_t$ ;
12   $\text{level}[i_t^*] = \max\{\text{level}[i_t^*], V_t\}$ ;
13  if  $X_t = 1$  then
14    sold[ $i_t^*$ ] = true;
15  end
16   $t = t + 1$ ;
17 end

```

with probability $\frac{1}{2}$. The customer's valuation is 2, so she will only buy the item if offered price 2, which occurs with total probability $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Note that:

1. Customer #6 increases the probability of unit 2 being sold from $\frac{1}{2}$ to $\frac{3}{4}$, which is consistent with her increasing $\text{level}[2]$ from 1 to 2;
2. Customer #6 increases the value of OPT , equal to $\sum_{i=1}^k \text{level}[i]$, by 1 (from 12 to 13);
3. Customer #6 brings in expected revenue $\frac{1}{4} \cdot 2 = \frac{1}{2}$.

Therefore, during time step 6, our procedure has earned expected revenue $\frac{1}{2} \cdot \Delta \text{OPT}$. We will show that it achieves this for a general \mathcal{P} , and all time steps t , regardless of the valuation of customer t .

2.2. Valuation Tracking Procedure and Analysis

We now formalize our valuation tracking procedure, in Algorithm 1.

In line 7, the procedure offers exactly one of the prices $r^{(\ell_t+1)}, \dots, r^{(m)}$, with the offering probabilities summing to unity. Note that it cannot branch to line 7 if $\ell_t = m$. This can be seen in the following way. If $\ell_t = m$, then i_t^* must have been assigned to some past customer t' with $V_{t'} = r^{(m)}$.

At time t' , either inventory unit i_t^* was already sold, or customer t' was offered a price at most $r^{(m)}$, which she would have accepted. In either case, $\text{sold}[i_t^*]$ must be **true**.

The analysis of Algorithm 1 is conceptually simple. Let $b_{i,t}$ be the index in $0, \dots, m$ such that $\text{level}[i] = r^{(b_{i,t})}$ at the end of time t , and j_t be the index such that $V_t = r^{(j_t)}$. We show that the following claims are maintained:

1. At the end of each time step t , the probability that any inventory unit i has been sold is

$$(\sum_{j=1}^{b_{i,t}} q^{(j)})/q;$$

2. During a time step t , if the valuation of the customer exceeds the level of the inventory unit she is assigned to, i.e. $j_t > \ell_t$, then:

(a) The expected revenue earned by Algorithm 1 is $\frac{1}{q}(r^{(j_t)} - r^{(\ell_t)})$;

(b) The increase in the offline optimum is $r^{(j_t)} - r^{(\ell_t)}$.

If $j_t \leq \ell_t$ during a time step t , then both the revenue of Algorithm 1 and the gain in OPT are 0.

These claims establish the following theorem, whose full proof is deferred to Section A.

THEOREM 1. *Algorithm 1 is $c^*(\mathcal{P})$ -competitive.*

2.3. Modified Algorithm based on Valuation Tracking Procedure

In this section we present a modified version of Algorithm 1 which is useful for the subsequent developments under the stochastic-valuation model in Section 3.

First we show how to modify Algorithm 1 so that its decision at each time t depends on only the remaining inventory, instead of the entire history of purchase decisions X_1, \dots, X_{t-1} .

DEFINITION 2. For all $t = 0, \dots, T$, let I_t denote the random variable for the amount of remaining inventory at the end of time t , which is equal to $k - \sum_{t'=1}^t X_{t'}$.

Our modified algorithm makes an offer to customer t according to the *probability* that unit i_t^* hasn't been sold, conditioned on the realized value of I_{t-1} . In this way, its decisions depend on only the inventory state, instead of the exact decisions of past customers.

DEFINITION 3 (ALGORITHM 1'). Define the following algorithm for choosing the price at each time t , based on the past valuations V_1, \dots, V_{t-1} and the amount of remaining inventory I_{t-1} .

1. Consider the indices i_t^* and ℓ_t during iteration t of Algorithm 1, which are deterministic based on V_1, \dots, V_{t-1} .
2. Compute the probability that $\text{sold}[i_t^*] = \text{true}$ on a run of Algorithm 1, *conditioned on* I_{t-1} units of inventory remaining after time $t-1$ in that run. Let γ_t denote this probability.
3. With probability γ_t , make an offer to customer t with the same price distribution as Algorithm 1 (line 7); with probability $1 - \gamma_t$, reject customer t .

Algorithm 1' chooses the distribution for P_t by “averaging” over all runs of Algorithm 1 which have the same value of I_{t-1} . We first remark that this can be done in polynomial time, despite there being exponentially many sample paths for Algorithm 1. We prove the following in Section B.

LEMMA 1. *The value of γ_t in Step 2 of Algorithm 1' can be computed in polynomial time.*

We now introduce some notation to disambiguate between random variables depicting the runs of different algorithms.

DEFINITION 4. For an algorithm \mathcal{A} , let $P_t^{\mathcal{A}}$, $X_t^{\mathcal{A}}$, and $I_t^{\mathcal{A}}$ be the random variables for the price at time t , purchase decision at time t , and inventory remaining at the end of time t , respectively. Let $\text{ALG}^{\mathcal{A}}$ be the random variable for the total revenue earned by algorithm \mathcal{A} . We will omit the superscripts \mathcal{A} when the context is clear.

Let $\mathcal{A} = \text{A1}$ refer to Algorithm 1 and $\mathcal{A} = \text{A1}'$ refer to Algorithm 1'.

We show that Algorithms 1 and 1' are virtually the same in that they have identical distributions for the remaining inventory at each time step, as well as the random price at each time step conditioned on any value of remaining inventory. This also establishes that Algorithm 1' is feasible, in that it does not try to make a sale with zero remaining inventory.

LEMMA 2. *For all $t \in [T]$, $k' \in \{0, \dots, k\}$ such that $\Pr[I_{t-1}^{\text{A1}'} = k'] > 0$, and $j \in \{1, \dots, m, m+1\}$, $\Pr[P_t^{\text{A1}'} = r^{(j)} | I_{t-1}^{\text{A1}'} = k'] = \Pr[P_t^{\text{A1}} = r^{(j)} | I_{t-1}^{\text{A1}} = k']$.*

Also, for all $t = 0, \dots, T$ and $k' \in \{0, \dots, k\}$, $\Pr[I_t^{\text{A1}'} = k'] = \Pr[I_t^{\text{A1}} = k']$.

Lemma 2, proven in Section B, is a consequence of the design of Algorithm 1'. For all t , the random price $P_t^{A1'}$ is identically distributed as P_t^{A1} , conditional on any value for the amount of remaining inventory at the end of time $t - 1$. Hence if $I_{t-1}^{A1'}$ and I_{t-1}^{A1} are identically distributed, then so are $I_t^{A1'}$ and I_t^{A1} . This allows us to inductively establish that the two algorithms have the same aggregate behavior after combining all sample paths, even though their behavior may differ given a specific history of purchase decisions. This also makes it easy to see that the expected revenues of the two algorithms are the same. Lemma 2 directly implies the following theorem.

THEOREM 2. *Algorithm 1' is $c^*(\mathcal{P})$ -competitive.*

2.4. Further Modified Algorithm and Structural Properties

In this section we present a further-modified version of Algorithm 1' which satisfies two structural properties: (i) it never rejects a customer if it has remaining inventory, offering the maximum price instead; (ii) the distribution of prices offered to a customer is *strictly stochastically-decreasing* (see Corollary 1) in the amount of remaining inventory. Property (ii) is the stochastic analogue of the classical structural result from Gallego and Van Ryzin (1994, Thm. 1) and its generalization to non-homogeneous demand in Zhao and Zheng (2000, Thm. 3): at any time step, if the firm has more inventory, then the optimal offering price is strictly lower.

DEFINITION 5 (ALGORITHM 1''). Define the following modification to Algorithm 1': in Step 3, offer price the maximum price $r^{(m)}$ to customer t , instead of rejecting her, with probability $1 - \gamma_t$.

We prove the following general lemma, which is intuitively easy to see, in Section B.

LEMMA 3. *Let \mathcal{A} be any pricing algorithm. Let \mathcal{A}' be the modified algorithm which: whenever \mathcal{A} would reject a customer while there is remaining inventory, \mathcal{A}' offers price $r^{(m)}$ instead. Then on any valuation sequence V_1, \dots, V_T , $\mathbb{E}[\text{ALG}^{\mathcal{A}'}] \geq \mathbb{E}[\text{ALG}^{\mathcal{A}}]$.*

Lemma 3 shows that Algorithm 1'' is $c^*(\mathcal{P})$ -competitive. Now, we would like to further show that the probability of Algorithm 1' rejecting, or correspondingly the probability of Algorithm 1'' offering the maximum price, is smaller when conditioned on larger values of remaining inventory.

THEOREM 3. *Suppose that the unconditional probability of Algorithm 1' rejecting customer t , $\Pr[\text{sold}[i_t^*] = \text{true}]$, lies in $(0, 1)$. Then for any $k_1 < k_2$ with $\Pr[I_{t-1} = k_1] > 0$ and $\Pr[I_{t-1} = k_2] > 0$,*

$$\Pr[\text{sold}[i_t^*] = \text{true} | I_{t-1} = k_1] > \Pr[\text{sold}[i_t^*] = \text{true} | I_{t-1} = k_2].$$

This structural property is intuitive, and we defer its proof to Section B. Theorem 3 allows us to conclude with the following corollary about strict stochastic dominance.

COROLLARY 1. *For any t , suppose $k_1 < k_2$ with $\Pr[I_{t-1} = k_1] > 0, \Pr[I_{t-1} = k_2] > 0$, and let $F_{P_t|I_{t-1}}(x|k_1), F_{P_t|I_{t-1}}(x|k_2)$ denote the CDF's of the price P_t offered by Algorithm 1" at time t conditional on the remaining inventory I_{t-1} being k_1, k_2 , respectively. Then $F_{P_t|I_{t-1}}(x|k_1) \leq F_{P_t|I_{t-1}}(x|k_2)$ for all $x \in \mathbb{R}$, and moreover, the dominance is strict in that $F_{P_t|I_{t-1}}(x|k_1) < F_{P_t|I_{t-1}}(x|k_2)$ for some price $x \in \mathcal{P}$.*

3. Stochastic Valuations

The model with stochastic valuations differs from the model with deterministic valuations studied in the previous section in the following ways. The valuation of each arriving customer is now *randomly* drawn from some probability distribution. The valuations of different customers are independent, but not necessarily identically distributed. An online algorithm is given the valuation *distribution* for each customer after the price for that customer has been chosen.

DEFINITION 6. We use the following notation, defined for all t :

- V_t : the valuation of customer t , a random variable taking values in $\{r^{(0)}, r^{(1)}, \dots, r^{(m)}\}$;
- \mathbf{v}_t : the probability vector $(v_t^{(0)}, v_t^{(1)}, \dots, v_t^{(m)})$ for the distribution of V_t , with $v_t^{(j)} = \Pr[V_t = r^{(j)}]$ and $\sum_{j=0}^m v_t^{(j)} = 1$;
- \mathbf{V}_t : (V_1, \dots, V_t) , the vector of realized valuations up to time t ;
- $\mathbf{P}_t^{\mathcal{A}}$: $(P_1^{\mathcal{A}}, \dots, P_t^{\mathcal{A}})$, the vector of prices up to time t chosen by algorithm \mathcal{A} .

We now provide a justification for our choice of offline optimum in our definition of competitiveness and competitive ratio, where in (2) we have replaced $\text{OPT}(V_1, \dots, V_T)$ with its expected value (and the values of V_t are realized independently).

3.1. Discussion of the Offline Optimum with Stochastic Valuations

The weakest (least clairvoyant) offline benchmark one could compare against with stochastic valuations is the following. Consider an offline algorithm which is given the sequence of valuation distributions, $\mathbf{v}_1, \dots, \mathbf{v}_T$, in advance. Given this sequence, it can solve for the policy which maximizes expected revenue, using dynamic programming. Clearly, the expected revenue of such a policy is an upper bound on the expected revenue obtainable by an online algorithm.

However, such a benchmark is difficult to compare against, because the optimal policy knowing $\mathbf{v}_1, \dots, \mathbf{v}_T$, while computable in polynomial time, may not admit any structure. Therefore, we relax the offline optimum by allowing the offline algorithm to know the *realizations* of all the valuations V_1, \dots, V_T in advance. The optimal algorithm knowing such information then has a trivial structure (sell to the k largest valuations).

In line with the definition from Section 2, let $\text{OPT}(V_1, \dots, V_T)$ denote the sum of the k largest valuations in V_1, \dots, V_T . We define the competitive ratio with stochastic valuations to be

$$\sup_{\text{ALG}} \inf_{\mathbf{v}_1, \dots, \mathbf{v}_T} \frac{\mathbb{E}[\text{ALG}(\mathbf{v}_1, \dots, \mathbf{v}_T)]}{\mathbb{E}_{V_1 \sim \mathbf{v}_1, \dots, V_T \sim \mathbf{v}_T}[\text{OPT}(V_1, \dots, V_T)]}. \quad (3)$$

In (3), the algorithm designer first commits to a (potentially randomized) algorithm ALG , after which an adversary chooses the sequence $\mathbf{v}_1, \dots, \mathbf{v}_T$, which determines the expected revenue of the online algorithm. The offline optimum is defined to be the expected value of $\text{OPT}(V_1, \dots, V_T)$ where V_1, \dots, V_T are realized according to $\mathbf{v}_1, \dots, \mathbf{v}_T$. We note that such an expected value cannot be computed in polynomial time; it is related to computing the expected project duration in a PERT network with independent task durations, which is $\#P$ -hard (Hagstrom 1988). Also, we should note that the definition of competitive ratio in (3) is equivalent to the following definition:

$$\sup_{\text{ALG}} \inf_{\mathbf{v}_1, \dots, \mathbf{v}_T} \mathbb{E} \left[\frac{\text{ALG}(\mathbf{v}_1, \dots, \mathbf{v}_T)}{\text{OPT}(V_1, \dots, V_T)} \right]$$

which is a general fact in competitive ratio analysis (Krumke 2002).

Finally, we should point out that a different relaxation in the offline optimum is possible, originating from Gallego and Van Ryzin (1994); see also Talluri and Van Ryzin (2006). Given the valuation distributions, $\mathbf{v}_1, \dots, \mathbf{v}_T$, one can write the following *Deterministic Linear Program (DLP)*:

$$\begin{aligned}
 \max \quad & \sum_{j=1}^m r^{(j)} \sum_{t=1}^T x_t^{(j)} \Pr[V_t \geq r^{(j)}] \\
 \sum_{j=1}^m \sum_{t=1}^T x_t^{(j)} \Pr[V_t \geq r^{(j)}] & \leq k \\
 \sum_{j=1}^m x_t^{(j)} & \leq 1 & t \in [T] \\
 x_t^{(j)} & \geq 0 & j \in [m], t \in [T]
 \end{aligned} \tag{4}$$

Let $\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T)$ denote the optimal objective value of the LP (4) where the value of each $\Pr[V_t \geq r^{(j)}]$ is computed based on \mathbf{v}_t . It can be shown that $\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T)$ is an upper bound on the expected revenue of the optimal dynamic programming policy knowing $\mathbf{v}_1, \dots, \mathbf{v}_T$, since $x_t^{(j)}$ encapsulates the unconditional probability of the policy offering price j to customer t .

Nonetheless, in this paper we don't compare against $\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T)$, which appears to be too strong of an offline benchmark. We show that the competitive ratio of $c^*(\mathcal{P})$, which was optimal in our deterministic setting, can still be achieved with stochastic valuations under definition (3). In other words, allowing the offline optimum to know the realizations of V_1, \dots, V_T does not decrease the competitive ratio. On the other hand, allowing the offline optimum to use the fractionality of the DLP (4) does decrease the competitive ratio; we provide some examples in Section E.

3.2. Optimally-Competitive Algorithm with Exponential Runtime

Having established our offline benchmark, we now derive $c^*(\mathcal{P})$ -competitive algorithms in the stochastic-valuation model. We do so by using our valuation tracking procedure as a subroutine, in a similar way to the development in Section 2.3, which may be helpful to reference.

Conceptually, our algorithm is a generalization of Algorithm 1' to stochastic valuations. However, since the assignment procedure in Algorithm 1 is no longer deterministic, we describe the algorithm in a different way. At a time step t , given $\mathbf{v}_1, \dots, \mathbf{v}_{t-1}$ and k' :

1. Consider a run of Algorithm 1 to the end of time t , where V_1, \dots, V_{t-1} are randomly drawn according to $\mathbf{v}_1, \dots, \mathbf{v}_{t-1}$. For all $j \in \{1, \dots, m, m+1\}$, compute $\Pr[P_t^{\text{A1}} = r^{(j)} | I_{t-1}^{\text{A1}} = k']$, where the probability is over both the random valuations and the random prices chosen by the algorithm. (If $\Pr[I_{t-1}^{\text{A1}} = k']$ has measure 0, then choose price $r^{(m+1)}$.)
2. For each $j \in \{1, \dots, m, m+1\}$, choose price $r^{(j)}$ with probability $\Pr[P_t^{\text{A1}} = r^{(j)} | I_{t-1}^{\text{A1}} = k']$.

Let Exp denote this algorithm, and we will use the corresponding notation from Definition 4.

It will be seen that Exp is a feasible policy when we establish that I_t^{Exp} and I_t^{A1} are identically distributed. First expand the expression $\Pr[P_t^{\text{A1}} = r^{(j)} | I_{t-1}^{\text{A1}} = k']$ as follows:

$$\frac{\Pr[P_t^{\text{A1}} = r^{(j)} \cap I_{t-1}^{\text{A1}} = k']}{\Pr[I_{t-1}^{\text{A1}} = k']} = \frac{\sum_{\mathbf{P}_{t-1}^{\text{A1}}, \mathbf{V}_{t-1}: I_{t-1}^{\text{A1}} = k'} \Pr[P_t^{\text{A1}} = r^{(j)} | \mathbf{P}_{t-1}^{\text{A1}}, \mathbf{V}_{t-1}] \cdot \Pr[\mathbf{P}_{t-1}^{\text{A1}} \cap \mathbf{V}_{t-1}]}{\sum_{\mathbf{P}_{t-1}^{\text{A1}}, \mathbf{V}_{t-1}: I_{t-1}^{\text{A1}} = k'} \Pr[\mathbf{P}_{t-1}^{\text{A1}} \cap \mathbf{V}_{t-1}]}. \quad (5)$$

In (5), the probability $\Pr[P_t^{\text{A1}} = r^{(j)} | \mathbf{P}_{t-1}^{\text{A1}}, \mathbf{V}_{t-1}]$, which conditions on a fixed history $P_1^{\text{A1}}, V_1, \dots, P_{t-1}^{\text{A1}}, V_{t-1}$, is defined by lines 6–10 of Algorithm 1. Thus, calculating (5) requires enumerating all histories that result in $I_{t-1}^{\text{A1}} = k'$.

Unfortunately, at each time step t , this takes time exponential in t . The computational difficulty arises because the assignment procedure in Algorithm 1 is no longer deterministic, as it was throughout Section 2. For now, we ignore computational constraints and focus on obtaining an $c^*(\mathcal{P})$ -competitive online algorithm; in Section 3.3 we show how to use sampling to achieve polynomial runtime while only losing ε in the competitiveness.

The following lemma is analogous to Lemma 2 and proved in Section C.

LEMMA 4. *For all $t \in [T]$, $k' \in \{0, \dots, k\}$ such that $\Pr[I_{t-1}^{\text{Exp}} = k'] > 0$, and $j \in \{1, \dots, m, m+1\}$,*

$$\Pr[P_t^{\text{Exp}} = r^{(j)} | I_{t-1}^{\text{Exp}} = k'] = \Pr[P_t^{\text{A1}} = r^{(j)} | I_{t-1}^{\text{A1}} = k']. \quad (6)$$

Also, for all $t = 0, \dots, T$ and $k' \in \{0, \dots, k\}$,

$$\Pr[I_t^{\text{Exp}} = k'] = \Pr[I_t^{\text{A1}} = k']. \quad (7)$$

Lemma 4 establishes that **Exp** is a feasible policy, i.e. it does not try to make a sale with no inventory remaining. Having established this, it remains to prove that **Exp** is optimally competitive. Theorem 4 is proved in Section C.

THEOREM 4. $\mathbb{E}[\text{ALG}^{\text{Exp}}(V_1, \dots, V_T)] = \mathbb{E}[\text{ALG}^{\text{A1}}(V_1, \dots, V_T)]$. By Theorem 1, $\text{ALG}^{\text{A1}}(V_1, \dots, V_T) = \frac{1}{q}\text{OPT}(V_1, \dots, V_T)$ for all realizations (V_1, \dots, V_T) . Therefore, **Exp** is $c^*(\mathcal{P})$ -competitive.

We should point out that although **Exp** does not inherit the polynomial-time property from Algorithm 1', it does inherit the structural property of the price at any time being stochastically-decreasing in the amount of remaining inventory. This is immediate from Theorem 3, which holds conditioned on any realization of V_1, \dots, V_T .

Also, note the following. P_t^{Exp} and P_t^{A1} are only guaranteed to be identically distributed when averaged over *all* the sample paths up to time $t - 1$ such that the total remaining inventory is k' . They may not be identically distributed when conditioned on a specific purchase sequence X_1, \dots, X_{t-1} such that $\sum_{t'=1}^{t-1} X_{t'} = k - k'$, or a specific valuation sequence V_1, \dots, V_{t-1} . Nonetheless, our method works in general. For example, if valuations were correlated, then we would condition on both I_{t-1} and V_1, \dots, V_{t-1} . One benefit of conditioning on only I_{t-1} in the independent case is to limit the state space, which is necessary for our polynomial-time sampling algorithm in Section 3.3.

3.3. Emulating the Exponential-runtime Algorithm using Sampling

In this section we show how to “emulate” **Exp**, using sampling, to achieve a polynomial runtime. First we provide a high-level overview of the challenges and the techniques used to overcome them. The intrinsic difficulty is that our original procedure is based on tracking the value of the offline optimum, but this becomes a #P-hard problem when the optimum equals the expected value of the k largest elements from independent realizations (see Hagstrom (1988)).

To overcome this using sampling, suppose we are at the start of time t , with inventory k' remaining. If we randomly sample a run of Algorithm 1 (drawing valuations randomly) such that $I_{t-1}^{\text{A1}} = k'$, and copy price P_t^{A1} for time t , then we would match the probabilities prescribed in (5).

This motivates the following algorithm: sample runs of Algorithm 1 to the end of $t - 1$ until hitting one where $I_{t-1}^{A1} = k'$, and then choose the price for time t according to lines 6–10 of Algorithm 1. Such an algorithm is equivalent to **Exp**, and thus would be $c^*(\mathcal{P})$ -competitive.

However, on sample paths where $\Pr[I_{t-1}^{A1} = k']$ is small, the sampling could take arbitrarily long. We limit the number of sampling tries so that the algorithm deterministically finishes in polynomial time, and show that the total measure of sample paths which fail at any point is $O(\varepsilon)$. Unfortunately, there could be correlation between the sampling failing, and having high revenue on a sample path. Nonetheless, we can *couple* the sample paths of the sampling algorithm to those of the exponential-time algorithm, mark the first point of failure on each sample path, and bound the difference in revenue after that point.

The details of the sampling algorithm, which we will call **Samp**, are specified in Algorithm 2.

In line 12, C is a positive integer to be chosen later. The decision of what to do when the sampling *fails*, i.e. defaults to line 13, is inconsequential, since in our analysis we do not expect any revenue from a sample path *after* the first point of failure.

To bound the revenue of Algorithm 2, we consider an algorithm which behaves identically to Algorithm 2, except even when it defaults to line 13, it is able to behave as if the sampling succeeded and makes the same decisions as lines 7–10. Such an algorithm is equivalent to **Exp**, and hereafter we will refer to it as **Exp**. The results of the sample runs do not affect the outcome of the algorithm, but help with bookkeeping.

DEFINITION 7. Let F_t^{Samp} be the indicator random variable for the sampling in Algorithm 2 failing at time t , defined for all $t \in [T + 1]$. Let $F_{T+1}^{\text{Samp}} = 1$ deterministically. Analogously, let F_t^{Exp} be the indicator random variable for the sampling in **Exp** “failing” at time t , $\forall t \in [T + 1]$.

For convenience, here we will use different random variables to denote the valuations in the runs of **Samp** and **Exp**: V_t^{Samp} and V_t^{Exp} , respectively. We will also use the notation from Definition 4.

DEFINITION 8. Define the *history up to time t* to consist of realizations up to and including the sampling at time t . Formally, for all $t \in [T + 1]$, let $\mathbf{h}_t = (f_1, p_1, v_1, \dots, f_{t-1}, p_{t-1}, v_{t-1}, f_t)$, where:

Algorithm 2: Weakly Randomized Online Algorithm based on Inventory Remaining

Input: Customers $t = 1, 2, \dots$ arriving online, with each valuation distribution \mathbf{v}_t revealed after the price P_t is chosen.

Output: For each customer t , a (possibly random) price P_t for her.

```

1 inventory =  $k$ ;
2  $t = 1$ ;
3 while customer  $t$  arrives do
4   repeat
5     run Algorithm 1 to the start of time  $t$ , with valuations  $V_1, \dots, V_{t-1}$  drawn according to
        $\mathbf{v}_1, \dots, \mathbf{v}_{t-1}$ , and prices  $P_1^{A1}, \dots, P_{t-1}^{A1}$  realized according to the random prices chosen
       by Algorithm 1;
6     if  $I_{t-1}^{A1} = \text{inventory}$  then
7       choose each price  $r^{(1)}, \dots, r^{(m)}, r^{(m+1)}$  according to the probability that Algorithm 1
       (on this run) would choose that price for customer  $t$ ;
8       observe  $\mathbf{v}_t$ ;
9       observe purchase decision of customer  $t$  and update inventory accordingly;
10       $t = t + 1$  and continue to next iteration of while loop;
11    end
12  until  $C(k+1)t^2$  runs elapse;
13  choose price  $\infty$ ;
14   $t = t + 1$ ;
15 end

```

- $f_{t'} \in \{0, 1\}$, for all $t' \in [t]$;
- $p_{t'}$ is a price in $\{r^{(1)}, \dots, r^{(m)}, r^{(m+1)}\}$, for all $t' \in [t-1]$;
- $v_{t'}$ is a valuation in $\{r^{(0)}, r^{(1)}, \dots, r^{(m)}\}$, for all $t' \in [t-1]$.

Furthermore, define the following vectors of random variables for all $t \in [T+1]$:

- $\mathbf{H}_t^{\text{Samp}} = (F_1^{\text{Samp}}, P_1^{\text{Samp}}, V_1^{\text{Samp}}, \dots, F_{t-1}^{\text{Samp}}, P_{t-1}^{\text{Samp}}, V_{t-1}^{\text{Samp}}, F_t^{\text{Samp}})$;
- $\mathbf{H}_t^{\text{Exp}} = (F_1^{\text{Exp}}, P_1^{\text{Exp}}, V_1^{\text{Exp}}, \dots, F_{t-1}^{\text{Exp}}, P_{t-1}^{\text{Exp}}, V_{t-1}^{\text{Exp}}, F_t^{\text{Exp}})$.

Now, we would like to partition the sample paths by the history up to the first point of failure, and prove that the two algorithms behave identically up to this point.

DEFINITION 9. Let \mathcal{F}_t denote the histories up to time t such that the first failure in the sampling occurs at time t . Formally, for all $t \in [T+1]$, \mathcal{F}_t is the set of $\mathbf{h}_t = (f_1, p_1, v_1, \dots, f_{t-1}, p_{t-1}, v_{t-1}, f_t)$ such that $f_1 = \dots = f_{t-1} = 0$ and $f_t = 1$. (p_1, \dots, p_{t-1} and v_1, \dots, v_{t-1} are arbitrary, and thus $|\mathcal{F}_t| = (m+1)^{2(t-1)}$.)

LEMMA 5. For a run of Algorithm 2, $\bigcup_{t=1}^{T+1} \bigcup_{\mathbf{h}_t \in \mathcal{F}_t} \{\mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t\}$ is a set of mutually exclusive and collectively exhaustive events. Analogously, for a run of Exp, $\bigcup_{t=1}^{T+1} \bigcup_{\mathbf{h}_t \in \mathcal{F}_t} \{\mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t\}$ is a set of mutually exclusive and collectively exhaustive events.

Furthermore, $\Pr[\mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t] = \Pr[\mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t]$, for all $t \in [T+1]$ and $\mathbf{h}_t \in \mathcal{F}_t$.

Lemma 5 is straight-forward, so we defer its proof to Section C. Having proved it, we can write:

$$\mathbb{E}[\text{ALG}^{\text{Samp}}] = \sum_{t=1}^{T+1} \sum_{\mathbf{h}_t \in \mathcal{F}_t} \mathbb{E}[\text{ALG}^{\text{Samp}} | \mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t] \Pr[\mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t] \quad (8)$$

$$\mathbb{E}[\text{ALG}^{\text{Exp}}] = \sum_{t=1}^{T+1} \sum_{\mathbf{h}_t \in \mathcal{F}_t} \mathbb{E}[\text{ALG}^{\text{Exp}} | \mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t] \Pr[\mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t]. \quad (9)$$

Since we also know that $\Pr[\mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t] = \Pr[\mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t]$, our goal is to compare the expected revenues of the two algorithms conditional on each history $\mathbf{h}_t \in \mathcal{F}_t$.

When $t = T+1$, i.e. the sampling never fails, it is easy to see that the two revenues are equal.

Indeed, for any $\mathbf{h}_{T+1} \in \mathcal{F}_{T+1}$:

$$\begin{aligned} \mathbb{E}[\text{ALG}^{\text{Samp}} | \mathbf{H}_{T+1}^{\text{Samp}} = \mathbf{h}_{T+1}] &= \mathbb{E}\left[\sum_{t=1}^T P_t^{\text{Samp}} \cdot \mathbb{1}(V_t^{\text{Samp}} \geq P_t^{\text{Samp}}) \middle| \mathbf{H}_{T+1}^{\text{Samp}} = \mathbf{h}_{T+1}\right] \\ &= \sum_{t=1}^T p_t \cdot \mathbb{1}(v_t \geq p_t) \\ &= \mathbb{E}[\text{ALG}^{\text{Exp}} | \mathbf{H}_{T+1}^{\text{Exp}} = \mathbf{h}_{T+1}]. \end{aligned} \quad (10)$$

LEMMA 6. Recall that $\mathbb{E}[\text{OPT}(V_1, \dots, V_T)]$ is the expected value of the offline optimum with V_1, \dots, V_T drawn independently according to $\mathbf{v}_1, \dots, \mathbf{v}_T$. For $t \leq T$ and $\mathbf{h}_t \in \mathcal{F}_t$,

$$\mathbb{E}[\text{ALG}^{\text{Exp}} | \mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t] - \mathbb{E}[\text{ALG}^{\text{Samp}} | \mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t] \leq \mathbb{E}[\text{OPT}(V_1, \dots, V_T)], \quad (11)$$

Proof. Consider any $t \in [T]$ and $\mathbf{h}_t \in \mathcal{F}_t$. We have

$$\begin{aligned} \mathbb{E}[\text{ALG}^{\text{Samp}} | \mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t] &\geq \mathbb{E}\left[\sum_{t'=1}^{t-1} P_{t'}^{\text{Samp}} \cdot \mathbb{1}(V_{t'}^{\text{Samp}} \geq P_{t'}^{\text{Samp}}) \middle| \mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t\right] \\ &= \sum_{t'=1}^{t-1} p_{t'} \cdot \mathbb{1}(v_{t'} \geq p_{t'}). \end{aligned} \quad (12)$$

Meanwhile, $\mathbb{E}[\text{ALG}^{\text{Exp}} | \mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t]$ can be decomposed into

$$\sum_{t'=1}^{t-1} p_{t'} \cdot \mathbb{1}(v_{t'} \geq p_{t'}) + \mathbb{E}\left[\sum_{t'=t}^T P_{t'}^{\text{Exp}} \cdot \mathbb{1}(V_{t'}^{\text{Exp}} \geq P_{t'}^{\text{Exp}}) \middle| \mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t\right]. \quad (13)$$

We elaborate on the second term in (13). Clearly, $\sum_{t'=t}^T P_t^{\text{Exp}} \cdot \mathbb{1}(V_t^{\text{Exp}} \geq P_t^{\text{Exp}})$ cannot exceed the sum of the $\min\{k, T - t + 1\}$ largest valuations to appear during t, \dots, T , which we denote by $\text{OPT}(V_t^{\text{Exp}}, \dots, V_T^{\text{Exp}})$. Furthermore, the random valuations $V_t^{\text{Exp}}, \dots, V_T^{\text{Exp}}$ are independent of the history $\mathbf{H}_t^{\text{Exp}}$ up to time t , so we can remove the conditioning and upper-bound (13) with

$$\sum_{t'=1}^{t-1} p_{t'} \cdot \mathbb{1}(v_{t'} \geq p_{t'}) + \mathbb{E}[\text{OPT}(V_t^{\text{Exp}}, \dots, V_T^{\text{Exp}})]. \quad (14)$$

The expectation in (14) is with respect to $V_t^{\text{Exp}}, \dots, V_T^{\text{Exp}}$ being drawn independently according to $\mathbf{v}_t, \dots, \mathbf{v}_T$. (14) in turn is no greater than $\sum_{t'=1}^{t-1} p_{t'} \cdot \mathbb{1}(v_{t'} \geq p_{t'}) + \mathbb{E}[\text{OPT}(V_1^{\text{Exp}}, \dots, V_T^{\text{Exp}})]$, where the random variables $V_1^{\text{Exp}}, \dots, V_{t-1}^{\text{Exp}}$ are *not* conditioned on the event $\mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t$. The proof of the lemma concludes by comparing this expression with (12). \square

Substituting (10), for $\mathbf{h}_{T+1} \in \mathcal{F}_{T+1}$, and (11), for $\mathbf{h}_1, \dots, \mathbf{h}_T \in \mathcal{F}_1, \dots, \mathcal{F}_T$, into (8) and (9), we conclude that

$$\mathbb{E}[\text{ALG}^{\text{Exp}}] - \mathbb{E}[\text{ALG}^{\text{Samp}}] \leq \mathbb{E}[\text{OPT}] \cdot \left(\sum_{t=1}^T \sum_{\mathbf{h}_t \in \mathcal{F}_t} \Pr[\mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t] \right). \quad (15)$$

By Definition 9, the expression in parentheses is the total probability of the sampling failing at any point, before choosing the final price P_T^{Samp} . We bound the term for each $t \in [T]$ separately. As t increases, the number of samples increases, so the probability of failure decreases:

LEMMA 7. *For all $t \in [T]$, $\sum_{\mathbf{h}_t \in \mathcal{F}_t} \Pr[\mathbf{H}_t^{\text{Samp}} = \mathbf{h}_t] \leq \frac{1}{eCt^2}$.*

Proof. Consider any $t \in [T]$. For all $\mathbf{h}_t \in \mathcal{F}_t$, let $G(\mathbf{h}_t) = (f_1, p_1, v_1, \dots, f_{t-1}, p_{t-1}, v_{t-1})$, which is the vector of the first $3(t-1)$ entries in \mathbf{h}_t . Let $\mathbf{G}_{t-1}^{\text{Exp}} = (F_1^{\text{Exp}}, P_1^{\text{Exp}}, V_1^{\text{Exp}}, \dots, F_{t-1}^{\text{Exp}}, P_{t-1}^{\text{Exp}}, V_{t-1}^{\text{Exp}})$, which is a vector of $3(t-1)$ random variables.

We can write $\sum_{\mathbf{h}_t \in \mathcal{F}_t} \Pr[\mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t]$ as

$$\sum_{\mathbf{h}_t \in \mathcal{F}_t} \Pr[F_t^{\text{Exp}} = 1 | \mathbf{G}_{t-1}^{\text{Exp}} = G(\mathbf{h}_t)] \Pr[\mathbf{G}_{t-1}^{\text{Exp}} = G(\mathbf{h}_t)]. \quad (16)$$

Now, for each $\mathbf{h}_t \in \mathcal{F}_t$, $\Pr[F_t^{\text{Exp}} = 1 | \mathbf{G}_{t-1}^{\text{Exp}} = G(\mathbf{h}_t)]$ is the probability that all $C(k+1)t^2$ independent runs of Algorithm 1 fail to match the inventory remaining at the start of time t according to \mathbf{h}_t . For convenience, define $I(\mathbf{h}_t) = k - \sum_{t'=1}^{t-1} \mathbb{1}(v_{t'} \geq p_{t'})$. Then

$$\Pr[F_t^{\text{Exp}} = 1 | \mathbf{G}_{t-1}^{\text{Exp}} = G(\mathbf{h}_t)] = (1 - \Pr[I_{t-1}^{\text{A1}} = I(\mathbf{h}_t)])^{C(k+1)t^2}, \quad (17)$$

where I_{t-1}^{A1} is the total inventory remaining at the start of time t in a run of Algorithm 1.

Therefore, we can partition the \mathbf{h}_t in \mathcal{F}_t by $I(\mathbf{h}_t)$. For all $k' \in \{0, \dots, k\}$, define $\rho_{t,k'} = \Pr[I_{t-1}^{A1} = k']$.

The following can be derived by substituting (17) into (16):

$$\begin{aligned} \sum_{\mathbf{h}_t \in \mathcal{F}_t} \Pr[\mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t] &= \sum_{k'=0}^k (1 - \rho_{t,k'})^{C(k+1)t^2} \sum_{\mathbf{h}_t \in \mathcal{F}_t: I(\mathbf{h}_t)=k'} \Pr[\mathbf{G}_{t-1}^{\text{Exp}} = G(\mathbf{h}_t)] \\ &\leq \sum_{k'=0}^k \exp(-\rho_{t,k'} C(k+1)t^2) \sum_{\mathbf{h}_t \in \mathcal{F}_t: I(\mathbf{h}_t)=k'} \Pr[\mathbf{G}_{t-1}^{\text{Exp}} = G(\mathbf{h}_t)]. \end{aligned} \quad (18)$$

At this point, we would like to argue that $\sum_{\mathbf{h}_t \in \mathcal{F}_t: I(\mathbf{h}_t)=k'} \Pr[\mathbf{G}_{t-1}^{\text{Exp}} = G(\mathbf{h}_t)] \leq \Pr[I_{t-1}^{\text{Exp}} = k']$. To see this, note that $\sum_{\mathbf{h}_t \in \mathcal{F}_t: I(\mathbf{h}_t)=k'} \Pr[\mathbf{G}_{t-1}^{\text{Exp}} = G(\mathbf{h}_t)] = \Pr[I_{t-1}^{\text{Exp}} = k' \cap (F_1^{\text{Exp}} = \dots = F_{t-1}^{\text{Exp}} = 0)]$.

Applying the second statement of Lemma 4, we see that $\Pr[I_{t-1}^{\text{Exp}} = k'] = \rho_{t,k'}$. Substituting into (18), the following can be derived:

$$\begin{aligned} \sum_{\mathbf{h}_t \in \mathcal{F}_t} \Pr[\mathbf{H}_t^{\text{Exp}} = \mathbf{h}_t] &\leq \sum_{k'=0}^k \rho_{t,k'} \exp(-\rho_{t,k'} C(k+1)t^2) \\ &\leq \sum_{k'=0}^k \frac{1}{C(k+1)t^2} \exp(-1) \\ &= \frac{1}{eCt^2}. \end{aligned}$$

The second inequality holds because for a single $\rho_{t,k'} \in [0, 1]$, the function $\rho_{t,k'} e^{-\rho_{t,k'} C(k+1)t^2}$ is maximized at $\rho_{t,k'} = \frac{1}{C(k+1)t^2}$. The proof of the lemma is now complete. \square

It now follows easily that the sampling algorithm is within ε of being optimally competitive.

THEOREM 5. *For all $\varepsilon > 0$, if we set $C = \lceil \frac{6}{e\pi^2\varepsilon} \rceil$ in line 12 of Algorithm 2, then it is $(\frac{1}{q} - \varepsilon)$ -competitive, and has runtime polynomial in $\frac{1}{\varepsilon}$, k , T , and m .*

Theorem 5 is straight-forward and proved in Section C.

4. No Information on Valuations

In this section we discuss whether it is possible for an online algorithm to be $c^*(\mathcal{P})$ -competitive without any information (before or after, deterministic or distributional) on the valuations.

First we show that this is impossible for any online algorithm which price-skims independently, i.e. realizes its random price at each time step using an independent source of random bits.

PROPOSITION 1. *Suppose that either: (i) $m \geq 2$ and valuations can be 0 (as usual); or (ii) $m \geq 3$ and valuations cannot be 0. (Recall that m is the number of prices.) Then for any online algorithm where each P_t chosen independently based on the sales history X_1, \dots, X_{t-1} , there exists a sequence V_1, \dots, V_T such that*

$$\frac{\mathbb{E}[\text{ALG}(V_1, \dots, V_T)]}{\text{OPT}(V_1, \dots, V_T)} < c^*(\mathcal{P}).$$

Proof of Proposition 1. Let the starting inventory $k = 1$.

First, it is easy to see that if the distribution of P_1 is not such that $\Pr[P_1 = r^{(j)}] = \frac{q^{(j)}}{q}$ for all $j \in [m]$, then for some deterministic instance consisting of a single valuation in $\{r^{(1)}, \dots, r^{(m)}\}$, $\frac{\mathbb{E}[\text{ALG}]}{\text{OPT}}$ will be strictly less than $\frac{1}{q}$. Therefore we can without loss of generality assume that $\Pr[P_1 = r^{(j)}] = \frac{q^{(j)}}{q}$ for all $j \in [m]$ (regardless of whether valuations can be 0).

Now suppose $m \geq 3$. Consider the distribution of P_2 conditioned on $X_1 = 0$. If $\Pr[P_2 \geq r^{(m)} | X_1 = 0] = 1$, then consider the instance $T = 2, V_1 = 1, V_2 = r^{(m-1)}$. $\text{OPT} = r^{(m-1)}$, which exceeds 1, since $m \geq 3$. Meanwhile, $q\mathbb{E}[\text{ALG}] = q\frac{1}{q} < \text{OPT}$. On the other hand, if $\Pr[P_2 \geq r^{(m)} | X_1 = 0] < 1$, then consider the instance $T = 3, V_1 = V_2 = r^{(m-1)}, V_3 = r^{(m)}$. $\text{OPT} = r^{(m)}$. $q\mathbb{E}[P_1 X_1] = r^{(m-1)}$, while $\mathbb{E}[X_1] = 1 - \frac{q^{(m)}}{q}$. The best case for the algorithm, given that $V_2 = r^{(m-1)}$, is $P_2 = r^{(m-1)}$ when $P_2 < r^{(m)}$. Let $\Pr[P_2 = r^{(m-1)} | X_1 = 0] = \alpha$, which we know is positive. In this case, $q\mathbb{E}[P_2 X_2] = r^{(m-1)}\alpha q^{(m)}$ and $\mathbb{E}[X_2] = \alpha \frac{q^{(m)}}{q}$. Hence $q\mathbb{E}[P_3 X_3]$ is at most $qr^{(m)}(1 - \mathbb{E}[X_1 + X_2]) = r^{(m)}(1 - \alpha)q^{(m)}$.

All in all, $q\mathbb{E}[\text{ALG}]$ is at most

$$\begin{aligned} r^{(m-1)} + r^{(m-1)}\alpha\left(1 - \frac{r^{(m-1)}}{r^{(m)}}\right) + r^{(m)}(1 - \alpha)\left(1 - \frac{r^{(m-1)}}{r^{(m)}}\right) &= r^{(m)} + \alpha\left(2r^{(m-1)} - \frac{(r^{(m-1)})^2}{r^{(m)}} - r^{(m)}\right) \\ &= r^{(m)} - \frac{\alpha}{r^{(m)}}(r^{(m)} - r^{(m-1)})^2 \end{aligned}$$

The term getting subtracted is non-zero since $\alpha > 0$ and $r^{(m)} > r^{(m-1)}$. Therefore, $q\mathbb{E}[\text{ALG}] < \text{OPT}$.

This completes the proof when $m \geq 3$, since $c^*(\mathcal{P}) = 1/q$.

The case where $m = 2$ and valuations can be 0 is argued analogously. If $\Pr[P_2 \geq r^{(2)}] = 1$, then consider the instance $T = 2, V_1 = 0, V_2 = r^{(1)}$. If $\Pr[P_2 < r^{(2)}] = 1$, then consider the instance $T = 3, V_1 = V_2 = r^{(1)}, V_3 = r^{(2)}$. In both cases, it can be seen that $q\mathbb{E}[\text{ALG}] < \text{OPT}$, completing the proof of Proposition 1. \square

However, we show that it is possible to be $c^*(\mathcal{P})$ -competitive if the online algorithm can price-skim in a “coordinated” fashion, with the same probabilities as in Eren and Maglaras (2010).

PROPOSITION 2. *Consider the following random-fixed-price policy:*

1. *Initially, choose a random price P which is equal to each $r^{(j)}$ with probability $q^{(j)}/q$;*
2. *Offer price P as long as there is remaining inventory.*

This policy is $c^(\mathcal{P})$ -competitive.*

Proof of Proposition 2. Consider any realization of the valuations, V_1, \dots, V_T . Iteratively define the following quantities, for j from m down to 1:

$$n^{(j)} = \min \left\{ \sum_{t=1}^T \mathbb{1}(V_t = r^{(j)}), k - \sum_{j'=j+1}^m n^{(j')} \right\}. \quad (19)$$

Essentially, for each j , $n^{(j)}$ denotes the number of valuations equal to $r^{(j)}$ that should be picked out when picking out the $\min\{k, T\}$ largest valuations. OPT is then equal to $\sum_{j=1}^m r^{(j)} n^{(j)}$.

Now consider the execution of the policy on this instance. For all $j \in [m]$, if the random fixed price P is equal to $r^{(j)}$, then the number of sales will be equal to $\min\{\sum_{t=1}^T \mathbb{1}(V_t \geq r^{(j)}), k\}$, which by definition is equal to $\sum_{j'=j}^m n^{(j')}$. Therefore,

$$\begin{aligned} \mathbb{E}[\text{ALG}] &= \frac{1}{q} \sum_{j=1}^m q^{(j)} r^{(j)} \sum_{j'=j}^m n^{(j')} \\ &= \frac{1}{q} \sum_{j'=1}^m n^{(j')} \sum_{j=1}^{j'} \left(1 - \frac{r^{(j-1)}}{r^{(j)}}\right) r^{(j)} \\ &= \frac{1}{q} \sum_{j'=1}^m n^{(j')} r^{(j')} \end{aligned}$$

which equals $\frac{1}{q}\text{OPT}$, completing the proof that the random fixed price is $c^*(\mathcal{P})$ -competitive. \square

It is known that correlated randomness is very powerful in the design of online algorithms (see, e.g., Karp et al. (1990), who derive an extremely elegant solution to the online matching problem using correlated randomness). Indeed, we can use our policy from Proposition 2 under our previous models with more information on the valuations and still have a $c^*(\mathcal{P})$ -competitive algorithm. However, this is impractical for several reasons. First, the fact that the random price must be fixed

makes it impossible to make use of additional information that may be available on the valuations (this issue was also raised in Eren and Maglaras (2010)). Second, the random-fixed-price policy does not show how the price should evolve as inventory is depleted, and does not satisfy the intuitive structural property in dynamic pricing that the price is greater if the remaining inventory is less (see Section 2.4). In Section 5, we show how our algorithms from Sections 2–3 can be adapted into the setting where the personal information of each customer can be used in determining her price.

5. Personalization Revenue Management Model

In this section we consider the personalized online revenue management setup introduced by Golrezaei et al. (2014), where:

- the stochastic decision of each customer can be modeled accurately upon her arrival to the e-commerce platform (by using her characteristics);
- however, the overall intensity and characteristics of customers to arrive over time is difficult to model (and treated as unknown/arbitrary).

In our model, this would correspond to the stochastic-valuation model in Section 3, with the change that the distribution of each V_t is given *before* the algorithm has to set a price, instead of after. The algorithms from Section 3 can still be applied, and will be $c^*(\mathcal{P})$ -competitive. Furthermore, it is not possible to be better than $c^*(\mathcal{P})$ -competitive even with this personalized information, as discussed in Section 1.1.

Nonetheless, in this section we specify how our online algorithms can exploit personalized information to strictly improve their decisions, while remaining $c^*(\mathcal{P})$ -competitive. Take any $c^*(\mathcal{P})$ -competitive algorithm \mathcal{A} for the stochastic-valuation model (e.g. the algorithm from Section 3.2, or a modification following Section 2.4 which never rejects customers before stocking out). For each time step t and inventory level $k' > 0$ such that $\Pr[I_{t-1}^{\mathcal{A}} = k'] > 0$, consider the distribution for the price P_t chosen by algorithm \mathcal{A} conditioned on $I_{t-1}^{\mathcal{A}} = k'$ (this depends on the previously-observed valuation distributions $\mathbf{v}_1, \dots, \mathbf{v}_{t-1}$). Since now we also know the distribution \mathbf{v}_t of valuation V_t , we can compute the probability of algorithm \mathcal{A} making a sale during time t ,

$$\sum_{j=1}^m \Pr[P_t^{\mathcal{A}} = r^{(j)} | I_{t-1}^{\mathcal{A}} = k'] \Pr[V_t \geq r^{(j)}], \quad (20)$$

as well as its expected revenue,

$$\sum_{j=1}^m r^{(j)} \Pr[P_t^{\mathcal{A}} = r^{(j)} | I_{t-1}^{\mathcal{A}} = k'] \Pr[V_t \geq r^{(j)}]. \quad (21)$$

We can interpret (21) as the reward given to the algorithm during time t in exchange for the probability (20) of consuming inventory. The price distribution used by the algorithm to obtain such an exchange was chosen without knowing the distribution of V_t . However, since now we do know the distribution of V_t , we can potentially make a decision which achieves *more* expected reward under the same consumption probability. Specifically, we solve the following LP:

$$\begin{aligned} \max \quad & \sum_{j=1}^m r^{(j)} \Pr[V_t \geq r^{(j)}] p_j(t, k') \\ \text{s.t.} \quad & \sum_{j=1}^m \Pr[V_t \geq r^{(j)}] p_j(t, k') = \sum_{j=1}^m \Pr[P_t^{\mathcal{A}} = r^{(j)} | I_{t-1}^{\mathcal{A}} = k'] \Pr[V_t \geq r^{(j)}] \\ & \sum_{j=1}^m p_j(t, k') \leq 1 \\ & p_j(t, k') \geq 0 \quad \forall j = 1, \dots, m \end{aligned} \quad (22)$$

$p_j(t, k')$ represents the probability that we should offer price j at time t , conditioned on the remaining inventory being k' . We know that setting each $p_j(t, k') = \Pr[P_t^{\mathcal{A}} = r^{(j)} | I_{t-1}^{\mathcal{A}} = k']$ is a feasible solution, and hence the optimal objective value of the optimization problem is at least (21). Let $\{p_j^*(t, k') : j = 1, \dots, m\}$ denote an optimal solution to the optimization problem, for all t and k' .

PROPOSITION 3. *Consider the online algorithm which, at each time step t , sets the price randomly according to probabilities $\{p_j^*(t, k') : j = 1, \dots, m\}$, where k' is the remaining inventory at the start of time t . Then for any sequence of valuation distributions $\mathbf{v}_1, \dots, \mathbf{v}_T$, the algorithm's total expected revenue is at least $\mathbb{E}_{V_1 \sim \mathbf{v}_1, \dots, V_T \sim \mathbf{v}_T}[\text{OPT}(V_1, \dots, V_T)]$.*

Proposition 3 is established in the same way as Theorems 2 and 4—for $t = 1, \dots, T$, we can inductively ensure from constraint (22) that the distribution for the starting inventory level matches that of \mathcal{A} . Since the algorithm has the same distribution for inventory state at each time t and earns at least as much revenue as \mathcal{A} in expectation on every inventory state, its total revenue must be at least the offline optimum (and in fact is often much better since it exploits personalization).

6. Conclusion

In this paper we have provided a general solution to the single-leg revenue management problem which yields the best-possible competitive ratio with limited demand information. Our policies unify the inventory-dependent booking policies in Ball and Queyranne (2009) with the random price-skimming policies in Eren and Maglaras (2010). An important feature of our policies is that they show at each time step how the *price distribution* should depend on inventory when the future is *unknown*, complementing classical results which show how the *optimal price* should depend on inventory when the future is *known*. Our policies were derived using a new “valuation tracking” technique, which geometrically tracks the optimum and hedges against the arrival sequence immediately ending *in the most inventory-conservative fashion*. We believe this to be of general interest for competitive ratio analysis.

Finally, we explain why our analysis of the pricing case, where each customer has a valuation and chooses the lowest fare not exceeding it, captures all *rational* choice models. Suppose instead that the firm could offer an *assortment* of fare classes, and that each customer has a ranked list of fare classes she is willing to purchase, and chooses the highest-ranked fare class that is offered to her. We can define V_t to be the maximum fare in the list that customer t is willing to purchase, and then the offline optimum would still be the k largest values from V_1, \dots, V_T . Meanwhile, we can modify the online algorithm so that whenever it would have offered price P_t , it now shows all fares greater than or equal to P_t . This algorithm would still make a sale whenever $V_t \geq P_t$, except now it has the opportunity to earn revenue greater than P_t , if customer t does not choose the lowest offered fare. As a result, our $c^*(\mathcal{P})$ -competitive algorithms under the pricing model imply corresponding $c^*(\mathcal{P})$ -competitive algorithms under the assortment model.

Nevertheless, we would like to end on two open questions related to the assortment generalization. First, our argument above assumes rational choice models; however in practice certain fare classes could be designed as “decoys” for other fare classes. Second, our algorithms imply an “assortment-skimming” distribution over revenue-ordered assortments, but this assumes there is no limit on the number of fare classes offered. We believe that assortment skimming under cardinality constraints is an interesting problem.

Acknowledgments

The authors would like to thank He Wang for insightful discussions.

References

- Araman VF, Caldentey R (2009) Dynamic pricing for nonperishable products with demand learning. *Operations research* 57(5):1169–1188.
- Araman VF, Caldentey R (2011) Revenue management with incomplete demand information. *Wiley Encyclopedia of Operations Research and Management Science* .
- Ball MO, Queyranne M (2009) Toward robust revenue management: Competitive analysis of online booking. *Operations Research* 57(4):950–963.
- Bergemann D, Schlag KH (2008) Pricing without priors. *Journal of the European Economic Association* 6(2-3):560–569.
- Besbes O, Zeevi A (2009) Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research* 57(6):1407–1420.
- Borodin A, El-Yaniv R (2005) *Online computation and competitive analysis* (cambridge university press).
- Chen L, Wu C (2016) Bayesian dynamic pricing with unknown customer willingness-to-pay and limited inventory. *Working Paper* .
- Chen X, Ma W, Simchi-Levi D, Xin L (2016) Dynamic recommendation at checkout under inventory constraint. *manuscript on SSRN* .
- Cheung WC, Ma W, Simchi-Levi D, Wang X (2018) Inventory balancing with online learning. *arXiv preprint arXiv:1810.05640* .
- den Boer AV (2015) Dynamic pricing and learning: historical origins, current research, and new directions. *Surveys in operations research and management science* 20(1):1–18.
- Eren SS, Maglaras C (2010) Monopoly pricing with limited demand information. *Journal of revenue and pricing management* 9(1-2):23–48.
- Gallego G, Van Ryzin G (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management science* 40(8):999–1020.

-
- Golrezaei N, Nazerzadeh H, Rusmevichientong P (2014) Real-time optimization of personalized assortments. *Management Science* 60(6):1532–1551.
- Hagstrom JN (1988) Computational complexity of pert problems. *Networks* 18(2):139–147.
- Karp RM, Vazirani UV, Vazirani VV (1990) An optimal algorithm for on-line bipartite matching. *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, 352–358 (ACM).
- Krumke SO (2002) Online optimization: Competitive analysis and beyond .
- Lan Y, Gao H, Ball MO, Karaesmen I (2008) Revenue management with limited demand information. *Management Science* 54(9):1594–1609.
- Lim AE, Shanthikumar JG (2007) Relative entropy, exponential utility, and robust dynamic pricing. *Operations Research* 55(2):198–214.
- Ma W, Simchi-Levi D (2017) Online resource allocation under arbitrary arrivals: Optimal algorithms and tight competitive ratios. *manuscript on SSRN* .
- Maglaras C, Meissner J (2006) Dynamic pricing strategies for multiproduct revenue management problems. *Manufacturing & Service Operations Management* 8(2):136–148.
- Talluri KT, Van Ryzin GJ (2006) *The theory and practice of revenue management*, volume 68 (Springer Science & Business Media).
- Yao ACC (1977) Probabilistic computations: Toward a unified measure of complexity. *Foundations of Computer Science, 1977., 18th Annual Symposium on*, 222–227 (IEEE).
- Zhang H, Shi C, Qin C, Hua C (2016) Stochastic regret minimization for revenue management problems with nonstationary demands. *Naval Research Logistics (NRL)* 63(6):433–448.
- Zhao W, Zheng YS (2000) Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Management science* 46(3):375–388.

E-Companion

Appendix A: Full Proof of Theorem 1

DEFINITION EC.1. Define the following:

- $S_{i,t}$: the indicator random variable for whether inventory unit i is sold by the end of time t , i.e. the value of `sold`[i] at the end of time t , defined for all $i \in [k]$ and $t = 0, \dots, T$;
- i_t : the inventory unit assigned to customer t , taking a value in $[k]$ for all $t \in [T]$;
- $\ell_{i,t}$: the value such that `level`[i] = $r^{(\ell_{i,t})}$ at the end of time t , taking a value in $\{0, 1, \dots, m\}$ for all $i \in [k]$ and $t = 0, \dots, T$;
- j_t : the value in $\{0, 1, \dots, m\}$ such that $V_t = r^{(j_t)}$, defined for all $t \in [T]$.

Fix the deterministic sequence of valuations V_1, \dots, V_T chosen by the adversary. i_t , $\ell_{i,t}$, and j_t are not random variables; they are determined by V_1, \dots, V_T .

We would like to write the random variables $S_{i,t}$ in terms of the other random variables. By definition, $S_{i,0} = 0$ for all $i \in [k]$. For $t > 0$, the following equations hold:

$$S_{i_t,t} = S_{i_t,t-1} + X_t; \quad (\text{EC.1})$$

$$S_{i,t} = S_{i,t-1}, \text{ for } i \neq i_t. \quad (\text{EC.2})$$

(EC.1)–(EC.2) are easy to see. In the algorithm, the only inventory unit that could potentially be sold during time t is i_t . This explains why (EC.2) holds for all $i \neq i_t$. It also explains why $S_{i_t,t} = 1$ if and only if $S_{i_t,t-1} = 1$ or $X_t = 1$. Furthermore, $S_{i_t,t-1}$ and X_t cannot both be 1, since the algorithm does not try to sell inventory unit i_t again at time t if it has already been sold. This completes the explanation for (EC.1).

We now analyze the state of the `sold` array during the execution of the algorithm.

LEMMA EC.1. *At the end of each time step t , the probability that any inventory unit i has been sold is $\frac{1}{q} \sum_{j=1}^{\ell_{i,t}} q^{(j)}$. Formally, for all $t = 0, \dots, T$,*

$$\mathbb{E}[S_{i,t}] = \frac{1}{q} \sum_{j=1}^{\ell_{i,t}} q^{(j)}, \text{ for } i \in [k]. \quad (\text{EC.3})$$

Proof. We proceed by induction on t . (EC.3) is true at time $t = 0$, where $\mathbb{E}[S_{i,0}] = 0$ and $\ell_{i,0} = 0$ for all $i \in [k]$.

Now suppose we are at the end of some time $t > 0$ and (EC.3) was true at the end of time $t - 1$. We need to prove that (EC.3) is still true at the end of time t . For $i \neq i_t$, $S_{i,t} = S_{i,t-1}$, by (EC.2). The value of `level`[i] is unchanged by the algorithm during time t , so $\ell_{i,t} = \ell_{i,t-1}$ as well. The inductive hypothesis from time $t - 1$ then establishes that $\mathbb{E}[S_{i,t}] = \frac{1}{q} \sum_{j=1}^{\ell_{i,t}} q^{(j)}$.

It remains prove $\mathbb{E}[S_{i_t,t}] = \frac{1}{q} \sum_{j=1}^{\ell_{i_t,t}} q^{(j)}$. This is immediate if j_t is no greater than $\ell_{i_t,t-1}$ (the value of the ℓ variable during iteration t of the algorithm), since both $S_{i_t,t}$ and $\ell_{i_t,t}$ would be unchanged. If $j_t > \ell_{i_t,t-1}$, the following can be derived (let $\ell = \ell_{i_t,t-1}$ for brevity):

$$\begin{aligned}
\mathbb{E}[S_{i_t,t}] &= \mathbb{E}[S_{i_t,t-1}] + \mathbb{E}[X_t] \\
&= \mathbb{E}[S_{i_t,t-1}] + \mathbb{E}[X_t | S_{i_t,t-1} = 0] \cdot \Pr[S_{i_t,t-1} = 0] \\
&= \frac{1}{q} \sum_{j=1}^{\ell} q^{(j)} + \Pr[X_t = 1 | S_{i_t,t-1} = 0] \left(1 - \frac{1}{q} \sum_{j=1}^{\ell} q^{(j)}\right) \\
&= \frac{1}{q} \sum_{j=1}^{\ell} q^{(j)} + \left(\sum_{j=\ell+1}^{j_t} \frac{q^{(j)}}{\sum_{j'=\ell+1}^m q^{(j')}}\right) \left(\frac{\sum_{j=\ell+1}^m q^{(j)}}{q}\right) \\
&= \frac{1}{q} \sum_{j=1}^{j_t} q^{(j)}.
\end{aligned}$$

The first equality follows from (EC.1) and the linearity of expectation. The second equality conditions on $S_{i_t,t-1}$ being 0, since the value of X_t is 0 if $S_{i_t,t-1} = 1$. The third equality uses the value of $\mathbb{E}[S_{i_t,t-1}]$ guaranteed by the inductive hypothesis. In the fourth equality, the probability of getting a sale, conditioned on Algorithm 1 reaching line 7, is equal to the probability of choosing a price at most $r^{(j_t)}$, the valuation of customer t . The final equality achieves the desired result because $j_t = \ell_{i_t,t}$, the new value for `level`[i] after line 12 of iteration t of the algorithm.

This completes the induction and the proof of the lemma. \square

Now we analyze the expected revenue of the algorithm, which is $\mathbb{E}[\text{ALG}]$, or $\sum_{t=1}^T \mathbb{E}[P_t X_t]$. As argued earlier, there cannot be a sale in a time step t where $j_t \leq \ell_{i_t,t-1}$, so for these time steps $X_t = 0$ and $\mathbb{E}[P_t X_t] = 0$. The following lemma derives the value of $\mathbb{E}[P_t X_t]$ when $j_t > \ell_{i_t,t-1}$.

LEMMA EC.2. *Suppose $j_t > \ell_{i_t,t-1}$ in a time step $t \in [T]$. Then the expected revenue earned by the algorithm during time step t is $\frac{1}{q}(r^{(j_t)} - r^{(\ell_{i_t,t-1})})$.*

Proof. Let $t \in [T]$ be any time step for which $j_t > \ell_{i_t,t-1}$. For brevity, let ℓ denote $\ell_{i_t,t-1}$. The following can be derived:

$$\begin{aligned}
\mathbb{E}[P_t X_t] &= \mathbb{E}[P_t X_t | S_{i_t,t-1} = 0] \Pr[S_{i_t,t-1} = 0] \\
&= \left(\sum_{j=\ell+1}^m r^{(j)} \mathbb{E}[X_t | P_t = r^{(j)}] \Pr[P_t = r^{(j)} | S_{i_t,t-1} = 0]\right) \left(1 - \Pr[S_{i_t,t-1} = 1]\right) \\
&= \left(\sum_{j=\ell+1}^m r^{(j)} \mathbb{1}[j_t \geq j] \Pr[P_t = r^{(j)} | S_{i_t,t-1} = 0]\right) \left(1 - \frac{1}{q} \sum_{j'=1}^{\ell} q^{(j')}\right) \\
&= \left(\sum_{j=\ell+1}^{j_t} r^{(j)} \frac{q^{(j)}}{\sum_{j'=\ell+1}^m q^{(j')}}\right) \left(\frac{\sum_{j'=\ell+1}^m q^{(j')}}{q}\right)
\end{aligned}$$

$$= \frac{1}{q} \sum_{j=\ell+1}^{j_t} r^{(j)} \left(1 - \frac{r^{(j-1)}}{r^{(j)}}\right)$$

The first equality conditions on $S_{i_t, t-1}$ being 0; note that $X_t = 0$ if $S_{i_t, t-1} = 1$. The second equality conditions on the value of P_t , where we drop the conditioning on $S_{i_t, t-1}$ in the term $\mathbb{E}[X_t | P_t = r^{(j)}]$ since $P_t \neq \infty$ already implies $S_{i_t, t-1} = 0$. This term becomes $\mathbb{1}[j_t \geq j]$ in the third equality, since it is deterministically 1 or 0 depending on whether $V_t \geq r^{(j)}$, or equivalently $j_t \geq j$. The third equality also uses Lemma EC.1, for the value of $\Pr[S_{i_t, t-1} = 1]$. The fourth equality uses the offering probabilities from line 7 of Algorithm 1. The fifth equality uses the explicit definition of $q^{(j)}$ from Definition 1, and it is easy to see that the final expression is equal to $\frac{1}{q}(r^{(j_t)} - r^{(\ell)})$. \square

Lemma EC.2 in turn implies the following lemma.

LEMMA EC.3. *The expected revenue earned by the algorithm up to time t , $\sum_{t'=1}^t \mathbb{E}[P_{t'} X_{t'}]$, is $\frac{1}{q} \sum_{i=1}^k r^{(\ell_{i,t})}$.*

Proof. The customers up to time t can be partitioned according to the inventory unit they were assigned, so

$$\sum_{t'=1}^t \mathbb{E}[P_{t'} X_{t'}] = \sum_{i=1}^k \sum_{t' \leq t: i_{t'}=i} \mathbb{E}[P_{t'} X_{t'}]. \quad (\text{EC.4})$$

Consider any i . For each t' assigned to i , $\mathbb{E}[P_{t'} X_{t'}]$ is 0 if $j_{t'} \leq \ell_{i, t'-1}$. Denote the remaining t' such that $j_{t'} > \ell_{i, t'-1}$ by t'_1, \dots, t'_N , where $N \geq 0$ and $t'_1 < \dots < t'_N$. Using Lemma EC.2,

$$\sum_{t' \leq t: i_{t'}=i} \mathbb{E}[P_{t'} X_{t'}] = \sum_{n=1}^N \frac{1}{q} (r^{(j_{t'_n})} - r^{(\ell_{i, t'_n-1})}).$$

Before time t'_n , $\text{level}[i]$ was last updated at time t'_{n-1} , so $\ell_{i, t'_n-1} = j_{t'_{n-1}}$. Therefore, the sum telescopes and the remaining term is $\frac{1}{q} r^{(j_{t'_N})}$ (note that $r^{(\ell_{i, t'_1-1})} = r^{(0)} = 0$). Now, $j_{t'_N} = \ell_{i, t'_N}$, and $\text{level}[i]$ is not updated again in time steps $t'_N + 1, \dots, t$, so $\ell_{i, t'_N} = \ell_{i, t}$. Substituting $\sum_{t' \leq t: i_{t'}=i} \mathbb{E}[P_{t'} X_{t'}] = \frac{1}{q} r^{(\ell_{i,t})}$ into (EC.4) completes the proof. \square

Having established the revenue of our online algorithm, we compare it to the offline optimum. Knowing the sequence of valuations V_1, \dots, V_T in advance, it is clear that the following algorithm is optimal:

1. Find the $\min\{k, T\}$ customers with the largest valuations;
2. Charge each of these customers t her maximum willingness-to-pay V_t ;
3. Reject all other customers.

The revenue OPT would be the the sum of the $\min\{k, T\}$ largest valuations.

DEFINITION EC.2. For all $t \in [T]$, let $M^k(t)$ be a vector consisting of the k largest elements from (V_1, \dots, V_t) , in any order. If $t < k$, fill in the remaining entries of $M^k(t)$ with zeros.

Then $\text{OPT} = \sum_{i=1}^k M_i^k(T)$, where $M_i^k(T)$ denotes the i 'th entry of $M^k(T)$. It turns out that $M^k(t)$ is closely tracked by the `level` array from Algorithm 1, as t progresses from 1 to T . Both $\ell_{i,t}$ (the value of `level`[i] at the end of time t) and $M^k(t)$ are deterministic functions of V_1, \dots, V_t .

LEMMA EC.4. *For all $t = 0, \dots, T$, the entries of the vector $(r^{(\ell_{1,t})}, \dots, r^{(\ell_{k,t})})$ is a permutation of the entries of the vector $M^k(t)$.*

Proof. We proceed by induction on t . At time $t = 0$, both $M^k(0)$ and $(r^{(\ell_{1,0})}, \dots, r^{(\ell_{k,0})})$ is a vector of k zeros, so the statement is true.

Now consider $t > 0$, and suppose that $M^k(t-1)$ is a permutation of $(r^{(\ell_{1,t-1})}, \dots, r^{(\ell_{k,t-1})})$. Therefore, a minimum entry in $M^k(t-1)$ is equal to a minimum entry in $(r^{(\ell_{1,t-1})}, \dots, r^{(\ell_{k,t-1})})$, which in turn is equal to $r^{(\ell_{i_t,t-1})}$, by Definition EC.1.

If $j_t > \ell_{i_t,t-1}$, or equivalently $V_t = r^{(j_t)} > r^{(\ell_{i_t,t-1})}$, then by the definition of $M^k(t)$, V_t must be added to $M^k(t-1)$ and replace any minimum entry equal to $r^{(\ell_{i_t,t-1})}$. Meanwhile, $\ell_{i_t,t} = j_t$, and $\ell_{i,t} = \ell_{i,t-1}$ for all $i \neq i_t$, thus the only change from $(r^{(\ell_{1,t-1})}, \dots, r^{(\ell_{k,t-1})})$ to $(r^{(\ell_{1,t})}, \dots, r^{(\ell_{k,t})})$ is that the entry at index i_t has been replaced by $r^{(j_t)}$. Since $M^k(t-1)$ and $(r^{(\ell_{1,t-1})}, \dots, r^{(\ell_{k,t-1})})$ go through the same change at time t , $M^k(t)$ is still a permutation of $(r^{(\ell_{1,t})}, \dots, r^{(\ell_{k,t})})$.

If instead $j_t \leq \ell_{i_t,t-1}$, then every entry of $M^k(t-1)$ is already at least $r^{(j_t)} = V_t$, so $M^k(t-1)$ incurs no change at time t . Similarly, $\ell_{i,t} = \ell_{i,t-1}$ for all $i \in [k]$, so $(r^{(\ell_{1,t-1})}, \dots, r^{(\ell_{k,t-1})})$ incurs no change as well. In both cases, we have established that $M^k(t)$ is a permutation of $(r^{(\ell_{1,t})}, \dots, r^{(\ell_{k,t})})$, completing the induction and the proof. \square

With Lemma EC.3 and Lemma EC.4, it is easy to establish the competitiveness of Algorithm 1. Our online algorithm is designed so that during each time step t , it earns exactly $\frac{1}{q}$ of the amount that the offline optimum would increase by with the addition of V_t , and it does not need to observe V_t beforehand to accomplish this.

Proof of Theorem 1. Fix any sequence of valuations (V_1, \dots, V_T) . $\mathbb{E}[\text{ALG}]$ is equal to $\sum_{t=1}^T \mathbb{E}[P_t X_t]$, which in turn is equal to $\frac{1}{q} \sum_{i=1}^k r^{(\ell_{i,T})}$, by Lemma EC.3. Meanwhile, $\text{OPT} = \sum_{i=1}^k M_i^k(T)$, and the entries of $M^k(T)$ is a permutation of the entries of $(r^{(\ell_{1,T})}, \dots, r^{(\ell_{k,T})})$, by Lemma EC.4. Therefore, $\text{OPT} = \sum_{i=1}^k r^{(\ell_{i,T})} = q \cdot \mathbb{E}[\text{ALG}]$, completing the proof of Theorem 1. \square

Appendix B: Proofs from Sections 2.3–2.4

Proof of Lemma 1. Let $i = i_t^*$, for brevity. Let $S_{i,t}$ be the indicator random variable for inventory unit i being sold by the end of time t . For all $k' \in \{0, \dots, k\}$,

$$\Pr[S_{i,t-1} = 0 | I_{t-1} = k'] = \frac{\Pr[I_{t-1} = k' | S_{i,t-1} = 0] \Pr[S_{i,t-1} = 0]}{\Pr[I_{t-1} = k' | S_{i,t-1} = 0] \Pr[S_{i,t-1} = 0] + \Pr[I_{t-1} = k' | S_{i,t-1} = 1] \Pr[S_{i,t-1} = 1]} \quad (\text{EC.5})$$

by Bayes' law. For all $t \in [T]$, $i \in [k]$, and $k' \in \{0, \dots, k\}$, we explain how to compute $\Pr[I_{t-1} = k' | S_{i,t-1} = 0]$ in polynomial time; $\Pr[I_{t-1} = k' | S_{i,t-1} = 1]$ can be computed analogously.

First we argue that the Bernoulli random variables $\{S_{i',t-1} : i' \in [k]\}$ are independent. To see this, note that the assignment procedure in Algorithm 1 is deterministic. Therefore, each $S_{i',t-1}$ is only dependent on the prices chosen for the customers assigned to i' , and while these prices could be dependent on each other, they are independent from the prices chosen for customers not assigned to i' .

Furthermore, $I_{t-1} = k - \sum_{i'=1}^k S_{i',t-1}$. By independence, $\Pr[I_{t-1} = k' | S_{i,t-1} = 0] = \Pr[\sum_{i' \neq i} S_{i',t-1} = k - k']$. $\sum_{i' \neq i} S_{i',t-1}$ is simply the sum of $k - 1$ independent Bernoulli random variables with known mean (from Lemma EC.1), hence the probability that it equals a specific value can be computed using dynamic programming.

We elaborate on the dynamic programming. For notational convenience, without loss of generality assume $i = k$. We will inductively for $a = 0, \dots, k - 1$ maintain the value of $\Pr[\sum_{i'=1}^a S_{i',t-1} = b]$ for all $b \in \{0, \dots, k\}$. It is easy to initialize this for $a = 0$. Given $\Pr[\sum_{i'=1}^a S_{i',t-1} = b]$ for all $b \in \{0, \dots, k\}$, note that

$$\Pr[\sum_{i'=1}^{a+1} S_{i',t-1} = b] = \Pr[\sum_{i'=1}^a S_{i',t-1} = b - 1] \Pr[S_{a+1,t-1} = 1] + \Pr[\sum_{i'=1}^a S_{i',t-1} = b] \Pr[S_{a+1,t-1} = 0]$$

for all $b \in \{0, \dots, k\}$. Each iteration of a can be computed in time linear in k , and there are less than k iterations.

$\Pr[I_{t-1} = k' | S_{i,t-1} = 1]$ can be computed analogously. It is clear that both procedures can be done in time $O(k^2)$ (ignoring the $O(t)$ time it may take to compute the assignment procedure), completing the proof of Lemma 1. \square

Proof of Lemma 2 and Theorem 2. We argue that Lemma 2 and Theorem 2 are the special cases of Lemma 4 and Theorem 4 from the stochastic-valuation model. It is easy to check that the statements are analogous, so it suffices to show that Exp (as defined in Section 3.2) executed on deterministic valuations is identical to Algorithm 1' (as defined in Section 2.3).

We show that the decision rule for a single time period t , and any amount of inventory remaining k' , is the same. Let $i = i_t^*$ and $\ell = \ell_t$, for brevity. Consider the values of i and ℓ during iteration t of Algorithm 1 (with the deterministic valuations V_1, \dots, V_T). First consider any $j = \ell + 1, \dots, m$.

$$\begin{aligned} \Pr[P_t^{A1} = r^{(j)} | I_{t-1}^{A1} = k'] &= \Pr[S_{i,t-1} = 0 | I_{t-1}^{A1} = k'] \cdot \frac{q^{(j)}}{\sum_{j'=\ell+1}^m q^{(j')}} \\ &= \Pr[S_{i,t-1} = 0 | I_{t-1}^{A1} = k'] \Pr[P_t^{A1} = r^{(j)} | S_{i,t-1} = 0, I_{t-1}^{A1} = k'] \\ &= \Pr[P_t^{A1} = r^{(j)} \cap S_{i,t-1} = 0 | I_{t-1}^{A1} = k'] \\ &= \Pr[P_t^{A1} = r^{(j)} | I_{t-1}^{A1} = k'] \end{aligned}$$

The first equality holds by the specification of algorithm Algorithm 1'. The second equality holds by the specification of Algorithm 1, where we can add the conditioning on $I_{t-1}^{A1} = k'$ in the second probability due to independence. The final equality follows because $P_t^{A1} = r^{(j)} \neq \infty$ implies $S_{i,t-1} = 0$.

If $j = m + 1$, then

$$\begin{aligned} \Pr[P_t^{A1'} = \infty | I_{t-1}^{A1'} = k'] &= \Pr[S_{i,t-1} = 1 | I_{t-1}^{A1} = k'] \\ &= \Pr[P_t^{A1} = \infty | I_{t-1}^{A1} = k'] \end{aligned}$$

since the event $S_{i,t-1} = 1$ occurs if and only if the event $P_t^{A1} = \infty$ occurs.

Finally, clearly if $j \leq \ell$, then both $\Pr[P_t^{A1'} = r^{(j)} | I_{t-1}^{A1'} = k']$ and $\Pr[P_t^{A1} = r^{(j)} | I_{t-1}^{A1} = k']$ are 0.

We have shown that $\Pr[P_t^{A1'} = r^{(j)} | I_{t-1}^{A1'} = k'] = \Pr[P_t^{A1} = r^{(j)} | I_{t-1}^{A1} = k']$ for all $j \in \{1, \dots, m, m+1\}$, so it is the same decision rule as **Exp**, completing the proof. \square

Proof of Lemma 3. Fix a valuation sequence V_1, \dots, V_T and consider any sample path in the execution of \mathcal{A} ; let the sample path be depicted by the sequence of random prices P_1^A, \dots, P_T^A . The revenue ALG^A on that sample path is given by $\sum_{t: V_t \geq P_t^A} P_t^A$; note that the cardinality of the set $\{t : V_t \geq P_t^A\}$ is at most k .

On that same sample path, the modified algorithm \mathcal{A}' would sell to the k customers with the smallest indices in $\{t : V_t \geq \min\{P_t^A, r^{(m)}\}\}$ (or all the customers in that set if its cardinality is less than k). Let \mathcal{S} denote the set of customers served by the modified algorithm. Let $b = |\{t \in \mathcal{S} : P_t^A = \infty\}|$, the number of customers with valuation $r^{(j)}$ served by the modified algorithm that were rejected by the original algorithm.

It is easy to see that

$$\begin{aligned} \text{ALG}^{A'} - \text{ALG}^A &= \sum_{t \in \mathcal{S}} \min\{P_t^A, r^{(m)}\} - \sum_{t: V_t \geq P_t^A} P_t^A \\ &\geq br^{(m)} - \sum_{t \in \mathcal{S}'} P_t^A \end{aligned} \tag{EC.6}$$

where \mathcal{S}' is the set of customers that are no longer served by \mathcal{A}' because it used up b extra units of inventory. Since $|\mathcal{S}'| \leq b$, and $P_t^A \leq r^{(m)}$ for all t such that $P_t^A \leq V_t$, it is immediate that (EC.6) is non-negative. Since this holds on every sample path for \mathcal{A} , we have completed the proof that $\mathbb{E}[\text{ALG}^{A'}] \geq \mathbb{E}[\text{ALG}^A]$. \square

Proof of Theorem 3 The inventory level I_{t-1} is equal to $k - \sum_{i=1}^k (1 - S_{i,t-1})$, where $S_{i,t-1}$ is the indicator random variable for inventory unit i being sold by the end of time $t - 1$. We will hereafter omit the subscript $t - 1$.

Each term $(1 - S_i)$ is independent and equal to 1 with probability $(\sum_{j=b_i+1}^m)q^{(j)}/q$, which is the probability that inventory unit i has not been sold. We will denote it using p_i and let $Y_i = 1 - S_i$,

for brevity. As long as b_i (the index in $0, \dots, m$ of the highest valuation assigned to inventory unit i) is not 0 or m , $p_i \in (0, 1)$. We will without loss of generality assume that $p_i \in (0, 1)$ for all i , redefining k and re-indexing as necessary (if $p_i = 0$ or $p_i = 1$ then Y_i is deterministic and we can remove it from analysis of the random sum). By the assumptions in the statement of the theorem, this re-indexing does not cause i_t^* to fall outside of $1, \dots, k$; in fact we can without loss of generality assume $i_t^* = 1$. Furthermore, k_1, k_2 are at least 0 at at most the re-defined k , since they correspond to inventory levels that are realized with non-zero probability.

After all of these transformations, the statement reduces to

$$\Pr[Y_1 = 1 | \sum_{i=1}^k Y_i = k_1] < \Pr[Y_1 = 1 | \sum_{i=1}^k Y_i = k_2] \quad (\text{EC.7})$$

where each Y_i is an independent Bernoulli random variable of probability $p_i \in (0, 1)$ and $0 \leq k_1 < k_2 \leq k$. Furthermore, we can without loss of generality assume that $k_2 = k_1 + 1$.

If $k_1 = 0$, then (EC.7) is clearly true, since the LHS is 0 while the RHS is non-zero. So assume that $k_1 > 0$ and we can rewrite (EC.7) as follows:

$$\begin{aligned} & \frac{\Pr[Y_1 = 1 \cap \sum_{i=1}^k Y_i = k_1]}{\Pr[\sum_{i=1}^k Y_i = k_1]} < \frac{\Pr[Y_1 = 1 \cap \sum_{i=1}^k Y_i = k_1 + 1]}{\Pr[\sum_{i=1}^k Y_i = k_1 + 1]} \\ & \frac{p_1 \Pr[\sum_{i=2}^k Y_i = k_1 - 1]}{p_1 \Pr[\sum_{i=2}^k Y_i = k_1 - 1] + (1 - p_1) \Pr[\sum_{i=2}^k Y_i = k_1]} < \frac{p_1 \Pr[\sum_{i=2}^k Y_i = k_1]}{p_1 \Pr[\sum_{i=2}^k Y_i = k_1] + (1 - p_1) \Pr[\sum_{i=2}^k Y_i = k_1 + 1]} \\ & \left(1 + \frac{(1 - p_1)}{p_1} \cdot \frac{\Pr[\sum_{i=2}^k Y_i = k_1]}{\Pr[\sum_{i=2}^k Y_i = k_1 - 1]} \right)^{-1} < \left(1 + \frac{(1 - p_1)}{p_1} \cdot \frac{\Pr[\sum_{i=2}^k Y_i = k_1 + 1]}{\Pr[\sum_{i=2}^k Y_i = k_1]} \right)^{-1} \end{aligned}$$

Therefore, it suffices to prove that:

$$\begin{aligned} & \frac{\Pr[\sum_{i=2}^k Y_i = k_1]}{\Pr[\sum_{i=2}^k Y_i = k_1 - 1]} > \frac{\Pr[\sum_{i=2}^k Y_i = k_1 + 1]}{\Pr[\sum_{i=2}^k Y_i = k_1]} \\ & \Pr[\sum_{i=2}^k Y_i = k_1]^2 > \Pr[\sum_{i=2}^k Y_i = k_1 + 1] \Pr[\sum_{i=2}^k Y_i = k_1 - 1] \\ & \left(\sum_{S \subseteq \{2, \dots, k\}: |S|=k_1} \prod_{i \in S} p_i \prod_{i \notin S} (1 - p_i) \right)^2 > \left(\sum_{S: |S|=k_1+1} \prod_{i \in S} p_i \prod_{i \notin S} (1 - p_i) \right) \left(\sum_{S: |S|=k_1-1} \prod_{i \in S} p_i \prod_{i \notin S} (1 - p_i) \right) \end{aligned} \quad (\text{EC.8})$$

After expanding, both sides are a sum of terms of the form

$$\prod_{i=2}^k p_i^{a_i} (1 - p_i)^{2-a_i} \quad (\text{EC.9})$$

where each a_i is 0, 1, or 2 and the sum $\sum_{i=2}^k a_i$ equals $2k_1$, the total number of times that a “positive” term p_i (as opposed to a “negative” term $(1 - p_i)$) appears in the product. Let b denote the total number of $i = 2, \dots, k$ such that $a_i = 1$, which must be even.

Now, observe that the total number of times the term (EC.9) appears in the LHS of the expansion of (EC.8) is $\binom{b}{b/2}$ (because we choose $b/2$ of the b indices that are “positive” to come from the first bracket; the remaining $b/2$ must come from the second bracket) while the total number of times this term appears in the RHS is $\binom{b}{b/2+1}$ (because we choose $b/2 + 1$ of the b indices that are “positive” to come from the first bracket), with the latter being strictly less. Furthermore, none of these terms are 0, since all of the values of p_i lie strictly between 0 and 1. Therefore, the inequality is strict, completing the proof of the theorem. \square

Appendix C: Proofs from Section 3

Proof of Lemma 4. We proceed by induction on t . (7) is true for $t = 0$, since $\Pr[I_0^{\text{Exp}} = k] = \Pr[I_0^{\text{A1}} = k] = 1$.

Now suppose $t > 0$ and that (7) has been established for time $t - 1$. Then for every k' such that $\Pr[I_{t-1}^{\text{Exp}} = k'] > 0$, (6) holds by definition. Indeed, since $\Pr[I_{t-1}^{\text{A1}} = k'] = \Pr[I_{t-1}^{\text{Exp}} = k']$ by the inductive hypothesis, $\Pr[I_{t-1}^{\text{A1}} = k'] > 0$ for such k' .

We now consider (7) for time t . Note that $I_t^{\text{Exp}} = I_{t-1}^{\text{Exp}} - \mathbb{1}(V_t \geq P_t^{\text{Exp}})$. Therefore,

$$\Pr[I_t^{\text{Exp}} = k'] = \Pr[I_{t-1}^{\text{Exp}} = k' + 1 \cap V_t \geq P_t^{\text{Exp}}] + \Pr[I_{t-1}^{\text{Exp}} = k' \cap V_t < P_t^{\text{Exp}}] \quad (\text{EC.10})$$

for $k' \in \{0, \dots, k - 1\}$, while

$$\Pr[I_t^{\text{Exp}} = k] = \Pr[I_{t-1}^{\text{Exp}} = k \cap V_t < P_t^{\text{Exp}}]. \quad (\text{EC.11})$$

Now, for any $k' \in \{0, \dots, k\}$, if $\Pr[I_{t-1}^{\text{Exp}} = k'] > 0$, then the following can be derived:

$$\begin{aligned} & \Pr[V_t \geq P_t^{\text{Exp}} | I_{t-1}^{\text{Exp}} = k'] \\ &= \sum_{j=1}^{m+1} \Pr[V_t \geq P_t^{\text{Exp}} | P_t^{\text{Exp}} = r^{(j)}, I_{t-1}^{\text{Exp}} = k'] \Pr[P_t^{\text{Exp}} = r^{(j)} | I_{t-1}^{\text{Exp}} = k'] \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^{m+1} \Pr[V_t \geq r^{(j)}] \Pr[P_t^{\text{Exp}} = r^{(j)} | I_{t-1}^{\text{Exp}} = k'] \\
&= \sum_{j=1}^{m+1} \Pr[V_t \geq r^{(j)}] \Pr[P_t^{\text{A1}} = r^{(j)} | I_{t-1}^{\text{A1}} = k'] \\
&= \Pr[V_t \geq P_t^{\text{A1}} | I_{t-1}^{\text{A1}} = k'].
\end{aligned} \tag{EC.12}$$

In the second equality, we remove the conditioning on $I_{t-1}^{\text{Exp}} = k'$, since the valuation V_t is an independent random variable unaffected by any history. The third equality follows because we have already established (6) for time t . The final equality also requires independence.

By the inductive hypothesis that (7) holds for time $t-1$, $\Pr[I_{t-1}^{\text{Exp}} = k'] = \Pr[I_{t-1}^{\text{A1}} = k']$. If $\Pr[I_{t-1}^{\text{Exp}} = k']$ is also non-zero, then the following can be derived using (EC.12):

$$\begin{aligned}
\Pr[V_t \geq P_t^{\text{Exp}} | I_{t-1}^{\text{Exp}} = k'] \Pr[I_{t-1}^{\text{Exp}} = k'] &= \Pr[V_t \geq P_t^{\text{A1}} | I_{t-1}^{\text{A1}} = k'] \Pr[I_{t-1}^{\text{A1}} = k'] \\
\Pr[V_t \geq P_t^{\text{Exp}} \cap I_{t-1}^{\text{Exp}} = k'] &= \Pr[V_t \geq P_t^{\text{A1}} \cap I_{t-1}^{\text{A1}} = k'].
\end{aligned} \tag{EC.13}$$

If instead $\Pr[I_{t-1}^{\text{Exp}} = k'] = \Pr[I_{t-1}^{\text{A1}} = k'] = 0$, then $\Pr[V_t \geq P_t^{\text{Exp}} \cap I_{t-1}^{\text{Exp}} = k'] \leq \Pr[I_{t-1}^{\text{Exp}} = k'] = 0$. Similarly, $\Pr[V_t \geq P_t^{\text{A1}} \cap I_{t-1}^{\text{A1}} = k'] = 0$, and therefore, (EC.13) still holds.

We can analogously to (EC.12) and (EC.13) derive for all $k' \in \{0, \dots, k\}$ that

$$\Pr[V_t < P_t^{\text{Exp}} \cap I_{t-1}^{\text{Exp}} = k'] = \Pr[V_t < P_t^{\text{A1}} \cap I_{t-1}^{\text{A1}} = k']. \tag{EC.14}$$

We can substitute (EC.13) and (EC.14) into (EC.10) to see that

$$\begin{aligned}
\Pr[I_t^{\text{Exp}} = k'] &= \Pr[I_{t-1}^{\text{A1}} = k' + 1 \cap V_t \geq P_t^{\text{A1}}] + \Pr[I_{t-1}^{\text{A1}} = k' \cap V_t < P_t^{\text{A1}}] \\
&= \Pr\left[\sum_{t'=1}^{t-1} X_{t'}^{\text{A1}} = k - k' - 1 \cap X_t^{\text{A1}} = 1\right] + \Pr\left[\sum_{t'=1}^{t-1} X_{t'}^{\text{A1}} = k - k' \cap X_t^{\text{A1}} = 0\right] \\
&= \Pr[I_t^{\text{A1}} = k']
\end{aligned}$$

for all $k' \in \{0, \dots, k-1\}$. We can similarly substitute (EC.14) into (EC.11) to see that $\Pr[I_t^{\text{Exp}} = k] = \Pr[I_t^{\text{A1}} = k]$. This completes the induction and the proof of Lemma 4. \square

Proof of Theorem 4. The following is straight-forward to derive:

$$\begin{aligned}
\mathbb{E}[\text{ALG}^{\text{Exp}}] &= \sum_{t=1}^T \mathbb{E}[P_t^{\text{Exp}} \cdot \mathbb{1}(V_t \geq P_t^{\text{Exp}})] \\
&= \sum_{t=1}^T \sum_{j=1}^{m+1} r^{(j)} \Pr[V_t \geq r^{(j)}] \Pr[P_t^{\text{Exp}} = r^{(j)}] \\
&= \sum_{t=1}^T \sum_{j=1}^{m+1} r^{(j)} \Pr[V_t \geq r^{(j)}] \sum_{k'=0}^k \Pr[P_t^{\text{Exp}} = r^{(j)} | I_{t-1}^{\text{Exp}} = k'] \Pr[I_{t-1}^{\text{Exp}} = k'] \\
&= \sum_{t=1}^T \sum_{j=1}^{m+1} r^{(j)} \Pr[V_t \geq r^{(j)}] \sum_{k'=0}^k \Pr[P_T^{\text{A1}} = r^{(j)} | I_{t-1}^{\text{A1}} = k'] \Pr[I_{t-1}^{\text{A1}} = k'] \\
&= \sum_{t=1}^T \sum_{j=1}^{m+1} r^{(j)} \Pr[V_t \geq r^{(j)}] \Pr[P_t^{\text{A1}} = r^{(j)}] \\
&= \sum_{t=1}^T \mathbb{E}[P_t^{\text{A1}} \cdot \mathbb{1}(V_t \geq P_t^{\text{A1}})].
\end{aligned}$$

The second and sixth equalities use the independence of V_t , while the fourth equality uses both statements of Lemma 4. The final expression is equal to $\mathbb{E}[\text{ALG}^{\text{A1}}]$, completing the proof of Theorem 4. \square

Proof of Lemma 5. The first statement is easy to see. Since every sample path fails at time $T+1$ by definition, for any sample path $\mathbf{H}_T^{\text{Samp}}$, it must have a unique *first* point of failure in $[T+1]$, say t' . $\mathbf{H}_T^{\text{Samp}}$ then falls under exactly one of the events, namely the one with $t = t'$ and $\mathbf{h}_t = (0, P_1^{\text{Samp}}, V_1^{\text{Samp}}, \dots, 0, P_{t-1}^{\text{Samp}}, V_{t-1}^{\text{Samp}}, 1)$. Therefore, the events are mutually exclusive and collectively exhaustive. The case for **Exp** is argued analogously.

The final statement is argued inductively. For all $t \in \{0, \dots, T\}$, let $\mathbf{g}_t = (f_1, p_1, v_1, \dots, f_t, p_t, v_t)$ be a vector of realizations to the end of time t , and let \mathcal{G}_t denote the set of such vectors containing no failures. Let $\mathbf{G}_t^{\text{Samp}} = (F_1^{\text{Samp}}, P_1^{\text{Samp}}, V_1^{\text{Samp}}, \dots, F_t^{\text{Samp}}, P_t^{\text{Samp}}, V_t^{\text{Samp}})$, and $\mathbf{G}_t^{\text{Exp}} = (F_1^{\text{Exp}}, P_1^{\text{Exp}}, V_1^{\text{Exp}}, \dots, F_t^{\text{Exp}}, P_t^{\text{Exp}}, V_t^{\text{Exp}})$.

We would like to inductively establish that $\Pr[\mathbf{G}_t^{\text{Samp}} = \mathbf{g}_t] = \Pr[\mathbf{G}_t^{\text{Exp}} = \mathbf{g}_t]$ for all $t \in \{0, \dots, T\}$ and $\mathbf{g}_t \in \mathcal{G}_t$. This is clearly true for $t=0$. For $t > 0$, take any $\mathbf{g}_t \in \mathcal{G}_t$, and we can write

$$\Pr[\mathbf{G}_t^{\text{Samp}} = \mathbf{g}_t] = \Pr[\mathbf{G}_{t-1}^{\text{Samp}} = \mathbf{g}_{t-1}] \cdot \Pr[F_t^{\text{Samp}} = 0 | \mathbf{G}_{t-1}^{\text{Samp}} = \mathbf{g}_{t-1}] \cdot \Pr[P_t^{\text{Samp}} = p_t | \mathbf{G}_{t-1}^{\text{Samp}} = \mathbf{g}_{t-1}, F_t^{\text{Samp}} = 0]$$

$$\begin{aligned}
& \cdot \Pr[V_t^{\text{Samp}} = v_t | \mathbf{G}_{t-1}^{\text{Samp}} = \mathbf{g}_{t-1}, F_t^{\text{Samp}} = 0, P_t^{\text{Samp}} = p_t]; \\
\Pr[\mathbf{G}_t^{\text{Exp}} = \mathbf{g}_t] &= \Pr[\mathbf{G}_{t-1}^{\text{Exp}} = \mathbf{g}_{t-1}] \cdot \Pr[F_t^{\text{Exp}} = 0 | \mathbf{G}_{t-1}^{\text{Exp}} = \mathbf{g}_{t-1}] \cdot \Pr[P_t^{\text{Exp}} = p_t | \mathbf{G}_{t-1}^{\text{Exp}} = \mathbf{g}_{t-1}, F_t^{\text{Exp}} = 0] \\
& \cdot \Pr[V_t^{\text{Exp}} = v_t | \mathbf{G}_{t-1}^{\text{Exp}} = \mathbf{g}_{t-1}, F_t^{\text{Exp}} = 0, P_t^{\text{Exp}} = p_t].
\end{aligned}$$

We will prove that $\Pr[\mathbf{G}_t^{\text{Samp}} = \mathbf{g}_t] = \Pr[\mathbf{G}_t^{\text{Exp}} = \mathbf{g}_t]$ by arguing that each term in the expression for $\Pr[\mathbf{G}_t^{\text{Samp}} = \mathbf{g}_t]$ equals the corresponding term in the expression for $\Pr[\mathbf{G}_t^{\text{Exp}} = \mathbf{g}_t]$. The first terms are equal because of the inductive hypothesis. The second terms are equal because both algorithms are sampling runs of Algorithm 1 and trying to hit a run with $I_{t-1}^{\text{A1}} = k - \sum_{t'=1}^{t-1} \mathbb{1}(v_{t'} \geq p_{t'})$. The third terms are identical because we have conditioned on $F_t^{\text{Samp}} = 0$. The fourth terms are equal because V_t^{Samp} and V_t^{Exp} are IID and none of the conditioning has any effect.

Having established this, note that for every $t \in [T+1]$ and $\mathbf{h}_t \in \mathcal{F}_t$ there exists a unique $\mathbf{g}_{t-1} \in \mathcal{G}_{t-1}$ such that \mathbf{g}_{t-1} is a prefix of \mathbf{h}_t . We know that for this \mathbf{g}_{t-1} , $\Pr[\mathbf{G}_{t-1}^{\text{Samp}} = \mathbf{g}_{t-1}] = \Pr[\mathbf{G}_{t-1}^{\text{Exp}} = \mathbf{g}_{t-1}]$. Therefore, it suffices to prove that $\Pr[F_t^{\text{Samp}} = 0 | \mathbf{G}_{t-1}^{\text{Samp}} = \mathbf{g}_{t-1}] = \Pr[F_t^{\text{Exp}} = 0 | \mathbf{G}_{t-1}^{\text{Exp}} = \mathbf{g}_{t-1}]$. By the same argument as the previous paragraph, these two probabilities are equal. Therefore, $\Pr[\mathbf{H}^{\text{Samp}} = \mathbf{h}_t] = \Pr[\mathbf{H}^{\text{Exp}} = \mathbf{h}_t]$, completing the proof of Lemma 5. \square

Proof of Theorem 5. Applying Lemma 7 to (15), we see that

$$\begin{aligned}
\mathbb{E}[\text{ALG}^{\text{Samp}}] &\geq \mathbb{E}[\text{ALG}^{\text{Exp}}] - \mathbb{E}[\text{OPT}] \left(\sum_{t=1}^T \frac{1}{e C t^2} \right) \\
&\geq \mathbb{E}[\text{ALG}^{\text{Exp}}] - \mathbb{E}[\text{OPT}] \frac{1}{e \lceil \frac{6}{e \pi^2 \varepsilon} \rceil} \left(\frac{\pi^2}{6} \right) \\
&\geq \mathbb{E}[\text{ALG}^{\text{Exp}}] - \varepsilon \cdot \mathbb{E}[\text{OPT}]
\end{aligned}$$

Furthermore, we know from Theorem 4 that $\mathbb{E}[\text{ALG}^{\text{Exp}}] = \mathbb{E}[\text{ALG}^{\text{A1}}] = \frac{1}{q} \mathbb{E}[\text{OPT}]$. This establishes the competitiveness.

The statement about runtime also follows easily from the specification of Algorithm 2 since the number of sample runs during each time period t , $\lceil \frac{6}{e \pi^2 \varepsilon} \rceil (k+1) t^2$, is polynomial in $\frac{1}{\varepsilon}$. \square

Algorithm 3: Weakly Randomized Online Algorithm for Continuum of Prices

Input: Customers $t = 1, 2, \dots$ arriving online, with each valuation V_t revealed after the price P_t is chosen.

Output: For each customer t , a (possibly random) price P_t for her.

```

1  val[i] = 0, sold[i] = false for  $i = 1, \dots, k$ ;
2   $t = 1$ ;
3  while customer  $t$  arrives do
4       $v = \min_{i'} \{\text{val}[i']\}$ ;
5       $i = \min\{i' : \text{val}[i'] = v\}$ ;
6      if sold[i] = false then
7          if  $v = 0$  then
8              offer price 1 w.p.  $\frac{1}{1+\ln R}$ , and price  $r$  w.p.  $\frac{1}{r(1+\ln R)}$  for all  $r \in (1, R]$ ;
9          else
10             offer price  $r$  w.p.  $\frac{1}{r(\ln R - \ln v)}$  for all  $r \in (v, R]$ ;
11         end
12     else
13         reject the customer by choosing price  $\infty$ ;
14     end
15     observe valuation  $V_t$  and purchase decision  $X_t$ ;
16     if  $V_t > v$  then
17         val[i] =  $V_t$ ;
18         if  $X_t = 1$  then
19             sold[i] = true;
20         end
21     end
22      $t = t + 1$ ;
23 end

```

Appendix D: A Continuum of Prices

In this section we show how to modify Algorithm 1 for the setting where valuations could take any value in $0 \cup [1, R]$. The competitive ratio obtained will be $\frac{1}{1+\ln R}$, recovering the competitive ratio from Ball and Queyranne (2009).

Consider Algorithm 3. Now $\text{val}[i]$ keeps track of the highest valuation assigned to inventory unit i thus far, starting at 0. It is easy to check that the price distributions specified in lines 8 and 10 are proper.

To analyze the competitiveness of Algorithm 3, we prove lemmas analogous to Lemmas EC.1–EC.2. We use the same notation as in Definition EC.1, except instead of $\ell_{i,t}$ and j_t , we use $v_{i,t}$ to denote the value of $\text{val}[i]$ at the end of time t , taking a value in $0 \cup [1, R]$.

LEMMA EC.5. *At the end of each time step t , the probability that any inventory unit i has been sold is 0 if $v_{i,t} = 0$, and $\frac{1+\ln v_{i,t}}{1+\ln R}$ if $v_{i,t} \geq 1$. Formally, for all $t = 0, \dots, T$,*

$$\mathbb{E}[S_{i,t}] = \mathbb{1}(v_{i,t} > 0) \cdot \frac{1 + \ln v_{i,t}}{1 + \ln R}, \text{ for } i \in [k]. \quad (\text{EC.15})$$

Proof. We proceed by induction on t . (EC.15) is true at time $t = 0$, where $\mathbb{E}[S_{i,0}] = 0$ and $v_{i,0} = 0$ for all $i \in [k]$.

Now suppose we are at the end of some time $t > 0$ and (EC.15) was true at the end of time $t - 1$. It suffices to prove that $\mathbb{E}[S_{i,t}] = \mathbb{1}(v_{i,t} > 0) \cdot \frac{1 + \ln v_{i,t}}{1 + \ln R}$. This is immediate if $V_t \leq v_{i,t-1}$, by the induction hypothesis. Otherwise, if $V_t > v_{i,t-1}$, we consider two cases. Let $v = v_{i,t-1}$ for brevity.

We know that $\mathbb{E}[S_{i,t}] = \mathbb{E}[S_{i,t-1}] + \mathbb{E}[X_t | S_{i,t-1} = 0] \cdot \Pr[S_{i,t-1} = 0]$.

If $v = 0$, then this equals

$$\begin{aligned} \Pr[X_t = 1 | S_{i,t-1} = 0] &= \frac{1}{1 + \ln R} \left(1 + \int_1^{v_{i,t}} \frac{1}{r} dr\right) \\ &= \frac{1 + \ln v_{i,t}}{1 + \ln R} \end{aligned}$$

as desired. On the other hand, if $v > 0$, then

$$\begin{aligned} \mathbb{E}[S_{i,t}] &= \frac{1 + \ln v}{1 + \ln R} + \frac{1}{\ln R - \ln v} \int_v^{v_{i,t}} \frac{1}{r} dr \left(1 - \frac{1 + \ln v}{1 + \ln R}\right) \\ &= \frac{1 + \ln v}{1 + \ln R} + \frac{\ln v_{i,t} - \ln v}{\ln R - \ln v} \left(\frac{\ln R - \ln v}{1 + \ln R}\right) \\ &= \frac{1 + \ln v_{i,t}}{1 + \ln R}. \end{aligned}$$

This completes the induction and the proof of the lemma. \square

LEMMA EC.6. *Suppose $V_t = v_{i,t} > v_{i,t-1}$ in a time step $t \in [T]$. Then the expected revenue earned by the algorithm during time step t is $\frac{1}{1 + \ln R}(v_{i,t} - v_{i,t-1})$.*

Proof. Let $t \in [T]$ be any time step for which $V_t > v_{i,t-1}$. Again, let v denote $v_{i,t-1}$, and we consider the two cases $v = 0$ and $v > 0$. If $v = 0$, then

$$\begin{aligned} \mathbb{E}[P_t X_t] &= \frac{1}{1 + \ln R} \left(1 + \int_1^R r \mathbb{E}[X_t | P_t = r] \frac{1}{r} dr\right) \\ &= \frac{1}{1 + \ln R} \left(1 + \int_1^R \mathbb{1}[v_{i,t} \geq r] dr\right) \\ &= \frac{1}{1 + \ln R} (1 + v_{i,t} - 1). \end{aligned}$$

In the first equality, a sale is guaranteed if $P_t = 1$, earning revenue 1. The final term is the desired expression.

If $v > 0$, then

$$\begin{aligned}\mathbb{E}[P_t X_t] &= \frac{1}{\ln R - \ln v} \left(\int_v^R r \mathbb{E}[X_t | P_t = r] \frac{1}{r} dr \right) \left(\frac{\ln R - \ln v}{1 + \ln R} \right) \\ &= \frac{1}{1 + \ln R} \left(\int_v^R \mathbb{1}[v_{i_t, t} \geq r] dr \right) \\ &= \frac{1}{1 + \ln R} (v_{i_t, t} - v)\end{aligned}$$

as desired. \square

With these two lemmas, the rest of the proof follows Section A. Indeed, Lemma EC.3 says that $\mathbb{E}[\text{ALG}] = \frac{1}{1 + \ln R} \sum_{i=1}^k v_{i, T}$. Meanwhile, Lemma EC.4 says that $\text{OPT} = \sum_{i=1}^k v_{i, T}$. Therefore, $\frac{\mathbb{E}[\text{ALG}]}{\text{OPT}} \geq \frac{1}{1 + \ln R}$, and since V_1, \dots, V_T was arbitrary, Algorithm 3 is $\frac{1}{1 + \ln R}$ -competitive.

Appendix E: Upper Bounds on the Competitive Ratio relative to the DLP

First, it is well-known that the DLP overestimates the optimum by a factor of $1 - \frac{1}{e}$, even when the feasible price set \mathcal{P} consists of a singleton (i.e. the dynamic pricing problem is trivial because there is only one price to choose from). The example requires the starting inventory k to be 1. Without loss of generality assume $\mathcal{P} = \{1\}$. Consider T customers, each of whom have a valuation exceeding 1 with probability $\frac{1}{T}$, and a valuation of 0 otherwise. It is easy to check that $\text{OPT}_{\text{LP}} = 1$ in this case, by setting $x_t^{(1)} = 1$ for all $t \in [T]$. Meanwhile, any algorithm cannot have expected revenue exceeding $1 - (1 - \frac{1}{T})^T$, where we have subtracted from 1 the probability of all customers having valuation 0. As $T \rightarrow \infty$, $\frac{\mathbb{E}[\text{ALG}]}{\text{OPT}_{\text{LP}}}$ approaches $1 - \frac{1}{e}$.

However, the gap becomes even larger if \mathcal{P} contains more than one price. We illustrate in the case of two feasible prices.

LEMMA EC.7. *Consider the stochastic-valuation model defined in Section 3, and let $\mathcal{P} = \{1, r\}$, $k = 1$. For all $r \geq 1$, there exists a distribution over $\mathbf{v}_1, \dots, \mathbf{v}_T$ such that for any online algorithm,*

$$\frac{\mathbb{E}[\text{ALG}(\mathbf{v}_1, \dots, \mathbf{v}_T)]}{\mathbb{E}[\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T)]} \leq \min\left\{1 - \frac{1}{e}, \frac{r - r/e}{2r - 1 - r/e}\right\}. \quad (\text{EC.16})$$

If $r \leq \frac{1}{1-1/e}$, then the RHS of (EC.16) is equal to $1 - \frac{1}{e} \approx .632$. However, if $r > \frac{1}{1-1/e}$, then we show that the upper bound is $\frac{r-r/e}{2r-1-r/e}$, which decreases to $\frac{e-1}{2e-1} \approx .387$ as $r \rightarrow \infty$.

Proof. Suppose that $r > \frac{1}{1-1/e}$, and let $p = \frac{1}{r(1-1/e)}$, which is in $(0, 1)$. Consider the following distribution over $\mathbf{v}_1, \dots, \mathbf{v}_T$:

- The first valuation distribution is deterministically $\mathbf{v}_1 = (v_1^{(0)}, v_1^{(1)}, v_1^{(2)}) = (0, 1, 0)$, i.e. the first customer deterministically has valuation 1.
- With probability p , valuation distributions $\mathbf{v}_2, \dots, \mathbf{v}_T$ are all equal to $(1 - \frac{1}{T-1}, 0, \frac{1}{T-1})$. When this occurs, each of the $T - 1$ customers $2, \dots, T$ are willing to pay r with probability $\frac{1}{T-1}$, and 0 otherwise.
- With probability $1 - p$, valuation distributions $\mathbf{v}_2, \dots, \mathbf{v}_T$ are all equal to $(1, 0, 0)$. When this occurs, all customers $2, \dots, T$ will never make a purchase.

We first compute the expected value of $\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T)$. With probability $1 - p$, $\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T) = 1$, setting $x_1^{(1)} = 1$. With probability p , $\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T) = r$, setting $x_1^{(1)} = x_1^{(2)} = 0$ and $x_2^{(2)} = \dots = x_T^{(2)} = 1$. Therefore, $\mathbb{E}[\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T)] = 1 - p + pr$.

We now consider the optimal strategy for the online algorithm. It has to decide, at time 1, whether to sell the only unit of inventory at price 1, without knowing whether $\mathbf{v}_2, \dots, \mathbf{v}_T$ are equal to $(1 - \frac{1}{T-1}, 0, \frac{1}{T-1})$ or $(1, 0, 0)$. Conditioned on it deciding to sell, $\text{ALG}(\mathbf{v}_1, \dots, \mathbf{v}_T)$ is deterministically 1. Conditioned on it deciding to wait, $\text{ALG}(\mathbf{v}_1, \dots, \mathbf{v}_T)$ is r with probability

$$\frac{1}{r(1-1/e)} \cdot (1 - (1 - \frac{1}{T-1})^{T-1}), \quad (\text{EC.17})$$

and 0 otherwise.

We explain (EC.17). If the online algorithm decides to wait, then it will offer price r to all customers beyond the first. It gets a sale if $\mathbf{v}_2 = \dots = \mathbf{v}_T = (1 - \frac{1}{T-1}, 0, \frac{1}{T-1})$, which occurs with probability $p = \frac{1}{r(1-1/e)}$, and further if at least 1 of the valuations V_2, \dots, V_T realizes to r , which yields the second term in (EC.17).

Thus the expected revenue from deciding to wait is (EC.17) multiplied by r , or

$$\frac{1}{(1-1/e)} \cdot (1 - (1 - \frac{1}{T-1})^{T-1}), \quad (\text{EC.18})$$

which is always greater than 1. Therefore, the online algorithm is better off waiting, in which case its expected revenue is (EC.18). Taking $T \rightarrow \infty$, (EC.18) approaches 1.

As $T \rightarrow \infty$, the distribution we constructed over $\mathbf{v}_1, \dots, \mathbf{v}_T$ is such that for the best online algorithm,

$$\begin{aligned} \frac{\mathbb{E}[\text{ALG}(\mathbf{v}_1, \dots, \mathbf{v}_T)]}{\mathbb{E}[\text{OPT}_{\text{LP}}(\mathbf{v}_1, \dots, \mathbf{v}_T)]} &= \frac{1}{1 - p + pr} \\ &= \frac{r(1 - 1/e)}{r(1 - 1/e) - 1 + r} \\ &= \frac{r - r/e}{2r - 1 - r/e}, \end{aligned}$$

as desired. \square