Certainty-equivalent pricing with dependent demand and limited price-changing opportunities

Hyun-Soo Ahn,* Christopher Thomas Ryan,† Joline Uichanco,‡ Mengzhenyu Zhang§

Abstract

When underlying demand follows a complex stochastic process, pricing problems are difficult to solve. In such cases, certainty equivalent (CE) policies, based on the solution to the deterministic relaxation of the stochastic pricing problem, can be used as practical alternatives. CE policies have lighter computational and informational requirements compared to solving the problem to optimality.

While the effectiveness of CE pricing policies has been theoretically studied when demands are independent, performance is not well-known when demands are state-dependent and price-changing opportunities are limited. This paper analyzes the performance of CE policies in a pricing problem where future demand depends on sales and inventory, and the firm has limited opportunities to change prices. We show that CE policies are asymptotically optimal: as the problem scale (denoted by m) becomes large, the percentage revenue loss decreases at the rate of $\Theta(1/\sqrt{m})$. We also extend the result to the joint pricing and (initial) inventory problem.

1 Introduction

In recent years, companies have used dynamic pricing as one of the levers to improve their sales revenue. First made popular in travel and hospitality industries with perishable inventory, dynamic pricing is now used in retail, logistics, services, and so on. The objective of dynamic pricing is to maximize the expected revenue over a finite selling horizon. An optimal dynamic pricing policy chooses the price that maximizes the expected revenue for the remainder of the horizon, given the current state (e.g., inventory, cumulative sales, etc.) and knowledge of demand.

In many settings, future demand is uncertain and depends on factors that can change over time. For example, when network effects drive future demand, then demand depends on cumulative sales. When inventory availability has a negative or positive effect on future demand (known as scarcity or display effects), then demand is affected by the state (e.g., inventory level) of the dynamical system.

For example, network effects are modeled in the seminal Bass diffusion model (Bass 1969), which has been shown to fit empirical demand curves of new products. In this model, sales of a

*University of Michigan Ross School of Business, hsahn@umich.edu
†UBC Sauder School of Business, chris.ryan@sauder.ubc.ca
‡University of Michigan Ross School of Business, jolineu@umich.edu
§University College London School of Management, ucznmzh@ucl.ac.uk

new product are primarily driven by word-of-mouth from previous customers. The display effect (i.e., demand is high when inventory is high) has been observed in sales data by Wolfe (1968), Smith and Achabal (1998) and Caro and Gallien (2012). This effect is attributed to more people noticing the product if there is more inventory. Conversely, the scarcity effect (i.e., demand is high when inventory is low or availability is limited) has been observed experimentally or empirically in Van Herpen et al. (2009), Balachander et al. (2009), Cui et al. (2019) and Cachon et al. (2018). This effect arises when the perceived value of a product increases when the item is exclusive or hard to get, creating a sense of urgency among customers to "act fast".

In addition to state-dependent demand, another consideration is that limited price change opportunities are predominant in practice. Because prices are only reviewed and changed at the start of each period, this setting is sometimes referred to as "periodic review" pricing (Yang and Zhang 2014; Bitran and Mondschein 1997). Many brick-and-mortar stores update their prices weekly, as changing prices often requires changing price stickers and POS data, which are costly and labor intensive. Although a continuous review of pricing is ubiquitous in analytical models of dynamic pricing, the ideal dynamic price cannot be implemented in many settings. However, fewer studies have analyzed optimal pricine policies in periodic review settings. There is a reason for this, as we will see, periodic pricing introduces technical challenges that require careful revision of existing analytical modes for solving dynamic pricing problems.

In addition to brick-and-mortar retail, technology and social networks have created new types of markets that did not exist before. These markets contain the feature of state-dependent demand and limited price changes. One example is a commerce market created by social media "influencers" who promote and sell products directly to their followers. A growing number of YouTube and Instagram "influencers" with a large subscriber base become sellers of their endorsed products or even their own products. This practice, broadly labeled as V-Commerce (virtual commerce), not only attracts interest in a product from the influencer's active followers but also provides convenient purchase options through the influencer's platforms. In V-commerce markets, the evolution of future demand follows a diffusion-like process based on hype and information propagation. Information about a product spreads as followers receive endorsements both from the influencer and from those in the influencer's network who have previously bought the product. Thus, not only is future demand dependent on the number of past adopters, but it can also be lumpy due to multiple customers buying the product simultaneously. Moreover, because the product is available for a relatively short period of time, the seller can change the price only a few times.

Certainty equivalent (CE) pricing policies are commonly used to analyze stochastic pricing problems like those studied in this paper (for example, Balseiro et al. (2023)). These policies rely on solving the deterministic counterparts of the stochastic problem by replacing all random variables with their expected values. An "open-loop" CE policy implements the optimal price sequence of the deterministic model. Although the actual prices of this policy can change during a sales season, they are static in the sense that the deterministic problem is solved once to obtain the price schedule for the entire season. In contrast, a "closed-loop" CE policy re-optimizes the deterministic model on a rolling horizon using the current inventory information at the beginning of each period. Hence, prices are adjusted over time based on the realizations of demands in past periods.

Both open-loop and closed-loop CE pricing policies are well-studied under a canonical setting where demands are independent across time, and price can be changed at any time (Gallego and Van Ryzin 1994; Jasin 2014). However, even though the phenomena of state-varying demand and limited price change opportunities are well-recognized to occur in practice, to the best of our knowledge, there has been yet no study of how CE pricing policies perform when the problem setting exhibits these features. For additional details on antecedant literature, see Table 1. Our work addresses this gap.

1.1 Technical challenges and contributions

Our setting presents two major technical challenges: (i) state-dependent demand leads to non-convex stochastic and deterministic problems, and (ii) limited-price changes which introduce the possibility of stockouts and demand censoring. The main contribution of our paper is to introduce a framework where it is nonetheless tractable to analyze the performance of certainty equivalent policies in this setting. Our framework is general enough so that it includes many of the state-dependent demand models proposed in the literature, such as Bass (1969); Datta and Pal (1990); Gerchak and Wang (1994); Urban and Baker (1997); Smith and Achabal (1998); Wang and Gerchak (2001); Shen et al. (2014); Smith and Agrawal (2017).

We quickly give an overview of how we achieve this. We start our analysis by establishing the tractability of solving for the optimal CE policies. This, at first, appears difficult since the deterministic version of the stochastic problem must grapple with demand censoring terms in the objective and non-convex constraints. However, through a series of transformations, we show that the problem is equivalent to a convex optimization model with a unique interior solution, which can be solved efficiently through interior point methods. Consequently, we show that solving for the CE policy is computationally tractable.

Next, we prove analytic performance for the CE policies, by bounding the gap between the CE expected revenues (under the unknown demand distribution) and the stochastic optimal expected revenue. We do this in two steps. First, we show that under any demand distribution whose conditional mean satisfies simple regularity assumptions, the optimal revenue of the deterministic model is an upper bound for the stochastic optimal expected revenue. However, we should note that standard techniques (i.e., Jensen's inequality, strong duality), which are used to establish a deterministic upper bound when demands are independent (e.g., Gallego and Van Ryzin 1994; Jasin 2014), cannot be used here due to non-convexity. Instead, we develop a novel induction argument to establish the upper bound through dynamic programming reformulations of the deterministic and stochastic pricing problems. This is new to the literature.

Then, we show that if the initial inventory and the expected demand are both scaled by the market size parameter, m, the gap between the expected revenue of a CE policy (both open-loop or closed-loop) and the deterministic upper bound grows in the order $\mathcal{O}(\sqrt{m})$, even when the price change opportunities are left fixed. We refer to this gap as the expected revenue loss. Since the deterministic revenue scales linearly in m, our analysis implies that both CE policies are asymptotically optimal as the problem scale increases.

For these bounding arguments, we cannot use standard techniques such as Scarf's bound, which are used to prove the $O(\sqrt{m})$ loss in the independent demand case. Instead, we show the $O(\sqrt{m})$ loss by carefully analyzing the demand censoring that might occur in each period by induction and showing that the sequences of states visited by the CE policies converge (as m increases) to the states visited by the deterministic optimal policy.

When demands are independent across periods, Jasin (2014) proved that re-optimization can reduce the revenue loss from $\mathcal{O}(\sqrt{m})$ to $\mathcal{O}(\log m)$ under the condition that more inventory strictly improves the revenue. Other researchers have since sharpened this bound to constant regret under certain conditions (Wang and Wang 2022; Jiang et al. 2022), but these improved arguments still remain tightly bound to the assumption of stationary demand and are hard to generalize to the dependent demand case. In particular, Jasin (2014) result requires explicit characterizations of the optimal primal and dual solutions of the re-optimized deterministic problem. Such an explicit characterization is not possible in our setting. In our general setting, we prove that the expected revenue loss of any policy that uses only expectation information is lower bounded by $\Omega(\sqrt{m})$. Hence, the $\mathcal{O}(\sqrt{m})$ bound on the expected revenue loss is tight for both open- and closed-loop CE policies. Exploring special settings in the dependent demand case where re-optimization can be shown to beat $\sqrt{m}$ loss is an open area for future research and would require generalizing the arguments of the existing literature (Jasin 2014; Wang and Wang 2022; Jiang et al. 2022) to a nonstationary setting or coming up with entirely new arguments.

The asymptotic regime considered in our setting is different from the asymptotic regimes considered in classical revenue management literature (e.g., Gallego and Van Ryzin 1994; Jasin 2014, etc.). In the literature, the number of price change opportunities always scales up in the scaled problem, allowing the opportunity for re-optimization policies to correct their pricing errors. However, in our asymptotic regime, when we scale up the demand rate and initial inventory, the number of price change opportunities is kept fixed; yet we are still able to show CE policies are asymptotically optimal despite much less pricing flexibility. Our setting of limiting the number of price changes (and the resulting need to deal with demand censoring) is a relevant and important extension because dynamic continuous price change cannot be applied to all situations in practice. For many products, such as electronics, computers, fashion products, and products sold through V-commerce, price changes occur a few times during the entire selling season. Thus, the limited number of price changes reflects these practical settings. We should note that our asymptotic results still hold if we scale up the number of price changes and initial inventory as well. (For a more detailed discussion of this and other technical issues, see Section 4.5).

We also show through large-scale numerical experiments that a small number of price changes is sufficient to recover nearly the same profit as an optimal policy for a continuous-time model with arbitrarily many price-change opportunities. The numerical experiments also provide guidance for choosing the number of price changes. In our simulations, as little as two to five price changes suffice to recover more than 95% of potential revenue from a continuous-time model. We also show that if the number of price change opportunities is less than the number of periods, the asymptotic performance of CE policies degrades to being linear in m. Together with our numerical study, this result shows that adding a little bit of price flexibility goes a long way.

We extend our analysis to the case where the firm needs to determine the initial inventory (in addition to prices) and show that the CE policy performs well in a joint pricing and inventory problem under state-dependent demand.

1.2 Literature review

In the operations literature, deterministic formulations are extensively studied, with a focus on deriving their structural properties. Thomas (1970); Rajan et al. (1992); Smith and Achabal (1998); Chen et al. (2001); Deng and Yano (2006); Geunes et al. (2006); Shen et al. (2014) study the joint decisions of pricing and production/inventory policies with deterministic demand. Sethi et al. (2008) propose the optimal advertising and pricing for a monopoly product under a deterministic demand process. Krishnamoorthy et al. (2010) extend the analysis to a duopoly market. Banker et al. (1998) use a deterministic optimization problem to study quality management. However, none of these papers theoretically analyze how well CE policies perform in stochastic settings.

On the other hand, the performance guarantee of CE policies are commonly studied in the revenue management literature, where such policies are adopted either because of their simplicity (Gallego and Van Ryzin 1994) or because the stochastic problem is difficult to solve (Gallego and Van Ryzin 1997; Bumpensanti and Wang 2020; Lei et al. 2021). Several papers establish theoretical performance bounds for CE policies. The vast majority of the papers that analyze CE policies for dynamic pricing problems make two general modeling assumptions (Gallego and Van Ryzin 1994, 1997; Jasin and Kumar 2012, 2013; Jasin 2014). First, demand is assumed to follow a specific stochastic process (e.g. a Poisson process) that depends only on the current price, so future demand is independent of the past demand. Second, they assume prices can be changed at any time. We refer to these two conditions as the *classical dynamic pricing setting.* The first condition results in a customer's purchase affecting the seller's current revenue but not its future demand. The second condition allows the seller to shut off demand immediately (by charging a high price) at the moment inventory runs out, implying that, almost surely, stockouts do not occur and demand censoring can be avoided. Together, these two conditions allow the associated CE problems to be formulated as linear or convex programs. Theoretical analyses of these settings directly utilize existing tools from linear or convex optimization (e.g., strong duality).

Under the assumption that customers arrive according to a homogeneous Poisson process, Gallego and Van Ryzin (1994) show that a fixed price is the solution to the CE problem, and the fixed-price CE policy is asymptotically optimal. In particular, they show the revenue loss of the CE pricing policy is $\mathcal{O}(\sqrt{m})$[1] when the total demand and the initial inventory are both scaled by m. Gallego and Van Ryzin (1994) is the first paper to show that, under certain conditions,

[1] Notation $\mathcal{O}, \Omega, \Theta$ are defined in Section 1.3.

Table 1: An overview of closely related papers in the literature.

	State-dependent demand	Limited price changes	Stockout	Inventory decision
Gallego and Van Ryzin (1994)	No	No	No	No
Bitran and Mondschein (1997)	No	Yes	Lost sales	Initial
Feng and Gallego (2000)	Yes	No[(a)]	No	No
Shen et al. (2014)	Yes	No	Backlog & lost sales	Replenish
Yang and Zhang (2014)	Yes	Yes	Backlog	Replenish
This paper	Yes	Yes	Lost sales	Initial

[(a)] Considers only finitely many price levels.

a fixed-price CE policy performs close to the optimal policy when price changes are possible for each arriving customer. Since then, a number of papers have shown similar guarantees for open-loop CE policies. For instance, Gallego and Van Ryzin (1997) and Jasin (2014) provide performance guarantees for open-loop CE controls in the network revenue management setting.

One potential weakness of an open-loop policy is that the price (which was computed assuming a representative sample path) is not adjusted to actual demand realizations. To overcome this, a number of papers examine the effectiveness of using re-optimization and modifying a CE policy with closed-loop feedback.

Some have studied settings in which closed-loop CE policies do not always outperform open-loop policies, such as in booking limit and bid price controls for network revenue management (Jasin and Kumar 2013). On the other hand, there are papers showing that closed-loop policies outperform open-loop policies (Maglaras and Meissner 2006; Chen and Farias 2013). Jasin and Kumar (2012) show that implementing a closed-loop CE policy in a probabilistic manner for a network revenue management (NRM) problem can have a revenue loss upper bounded by $\mathcal{O}(1)$, which is independent of the problem scale. Bumpensanti and Wang (2020) establish a similar loss bound by re-solving the deterministic linear program approximation for the NRM problem under a less restrictive assumption. Reiman and Wang (2008) propose a closed-loop CE pricing policy where the re-solving time is endogenously determined by a heuristic. The expected revenue loss of their policy is $o(\sqrt{m})$.

We conclude this section with a table (Table 1) that positions our paper among those we found closest to our setting. As the reader can see, antecedent models in the dynamic pricing literature share some (but not all) of the features of our framework. The dynamic pricing literature is vast, each paper in the table is only representative of a number of papers with related questions, models, and results.

Recently, a number of other papers that touch on similar themes have appeared in the literature. We briefly describe the most relevant papers here and our relationship to them.

Balseiro et al. (2023) look at limits in the effectiveness of certainty equivalent fluid approximations. The paper provides an unifying framework in the context of the a Dynamic Resource-

Constrained Reward Collection (DRC^2) problem and attempt to "standardize" the usual asymptotic optimality found in several papers. All problems that are cast as a DRC^2 problems assume continuous-time and randomness in demand to be independent across periods. With this set-up, they show the performance of a fluid certainty-equivalent control heuristic in DRC^2 problems. By contrast, our problem *does not* fit to a DRC^2 problem. In our problem, demand is not necessarily independent, and whereas DRC^2 problems are in continuous time (or discrete time with at most one arrival per period), our problem has discrete time that allows multiple arrivals per period, which results in demand censoring. Moreover, the open-loop certainty equivalent heuristic (CE-OL) is not considered in Balseiro et al. (2023).

Next, we discuss two related papers: Bai et al. (2023) and Aouad and Ma (2022). A major part of their motivation was to also move beyond the "standard" arguments represented in the compelling distillation of Balseiro et al. (2023). Both papers, in particular, study the settings where the coefficient of variation of the number of customer arrivals exceeds $1/\sqrt{T}$ (which is analogous to our Assumption 3).

Bai et al. (2023) consider a random selling horizon model to allow for "high-variance demand" in the setting of a resource allocation problem. The naive fluid approximation of their problem is a linear program, which fails to approximate the optimal total expected revenue as the resource capacities get large. As a result, they propose an alternative linear fluid approximation with a modified constraint, and this reformulated fluid approximation is asymptotically optimal. The linear fluid approximation was possible because customer arrivals are state-independent, and the seller can make decisions after each arrival. On the other hand, we examine the problem with limited price changes and dependent demand. As a result, the fluid approximation of our problem is not a linear program, thus the arguments used in their paper no longer apply.

Turning to Aouad and Ma (2022), they study an online matching problem that untethers variance from being tightly bound to the mean. In their paper, there is no time notion—the queries/arrivals occur in one period and the queries of each type are independent of each other. As a result, the dependence of demands across periods is not captured in their model.

Li et al. (2023) considers the same problem studied in Aouad and Ma (2022); Bai et al. (2023) but allows customer arrivals in successive periods to be dependent. From this perspective, this paper is the most related to ours. However, a few key differences exist. Li et al. (2023) is not a dynamic pricing paper. Instead, they study dynamic capacity control where the decision is whether to "accept" or "reject" each arriving customer. On the other hand, our paper considers price as the decision and the price at the current period not only affects the current period demand but also future demand. From an analysis perspective, the fluid approximation in Li et al. (2023) is still a linear program. The linear form of the fluid approximation allows several analytical conveniences: the expected approximated revenue can be written as a simple form and there exists a closed linear form of the relaxed dynamic program (Lagrangian). Whereas, in our setting, the nonlinearity of the fluid approximation makes their analysis invalid. Jiang (2023) is a relevant paper in line with Li et al. (2023). Their paper uses a linear program to approximate the dynamic program by assuming the value-to-go function is linear in the remaining capacities.

Our problem is fundamentally different because we consider the pricing control, and the pricing decision affects the transition probabilities. *Thus, applying the same LP approximation of our dynamic program would generate an infinite number of constraints.* However, in Jiang (2023), the decision of accepting or not does not affect the transition probability of the system. The correlated customer arrivals are reflected in the transition probability depending on the customer type that arrived in the previous period.

Limited price experimentation is usually considered in the demand learning and pricing literature. For example, Cheung et al. (2017) considers limited price change opportunities in an independent demand setting, and Zhang et al. (2022) considers a specific state-dependent demand setting. However, in this paper, the length between price changes was not chosen strategically as Zhang et al. (2022) did. The pricing policies considered are commonly referred to as periodic pricing policies. We focus on the value of using pricing policies based on the certainty equivalent problem when demand at a time is not necessarily independent from the past demand. Thus, we contribute to the literature by extending the conditions in which solutions from deterministic problems become high-quality solutions for stochastic optimization problems. Our paper builds the foundation of the known demand case in the dependent demand and periodic pricing setting, and thus can serve as a foundation for later study that incorporates demand learning.

1.3 Preliminaries

Throughout the paper, we use the big $\mathcal{O}$ notation in expressions $f(x) = \mathcal{O}(g(x))$ where f and g are positive real-valued functions if there exists an $r \in \mathbb{R}$ such that $f(x) < rg(x)$ for x sufficiently large. Similarly, if $f(x) = \Omega(g(x))$, then $f(x) > rg(x)$. When $f(x) = \mathcal{O}(g(x))$ and $f(x) = \Omega(g(x))$, it is represented by $f(x) = \Theta(g(x))$.

2 Modeling framework

A monopolist is selling a product with finite inventory $\alpha > 0$ over a finite horizon. The firm can dynamically change the price, but these price changes can only occur at certain price review periods $\{1, 2, \ldots, T\}$, which are equally spaced during the selling horizon. Thus, T also denotes the number of prices. After the firm chooses a price $\pi_t \geq 0$ for period t, a random variable $D_t \geq 0$ is realized, representing the demand in period t. After the demand D_t is realized, it is satisfied to the maximum extent using the remaining inventory. We denote the remaining inventory at the end of period t as N_t, where $N_0 = \alpha$. Any unmet demand is lost. Goods not sold by the end of period T are salvaged at a (normalized) value of 0. We assume that conditional on the state at period t and the price, the expectation of D_t is a known function of the price π_t, of the cumulative past sales, and of the remaining inventory.

We present pricing policies that only make use of information on the conditional expectation of demand and analyze their performance in an asymptotic setting (specified in Section 4.2). In the asymptotic setting, we scale both the expected demand rate and the initial inventory by a factor $m > 0$ while keeping the number of price changes T fixed.

Notation	Description
T	number of price review periods
π_t	price at period t
D_t	stochastic demand in period t
N_t	remaining inventory at the end of period t
α	initial inventory level
$x(\pi_t)$	price sensitivity function of demand
$\lambda(N_{t-1}, \alpha)$	sales and inventory sensitivity (SIS) function of demand

Table 2: Notation for modeling framework.

2.1 Demand model

Let P_t denote the total cumulative demand up to period t, where $P_t = \sum_{s=1}^{t} D_s$. We define $\mathcal{F}_t = \sigma(P_0, P_1, \ldots, P_t)$ to be the smallest σ-field where variables $P_0, P_1, \ldots, P_t$ are measurable and let $\mathcal{F} = \{\mathcal{F}_0, \mathcal{F}_1, \ldots\}$ be the associated filtration. We assume that conditional on $\mathcal{F}_{t-1}$ and the price π_t, the distribution of D_t depends on the price, on the cumulative sales $\alpha - N_{t-1}$, and on the remaining inventory N_{t-1}. Note that the cumulative sales $\alpha - N_{t-1}$ is not the same as the cumulative demand P_{t-1}. It is possible that $\alpha - N_{t-1} < P_{t-1}$, which happens whenever the seller stocks out due to the cumulative demand P_{t-1} exceeding the initial inventory $N_0 = \alpha$.

Assumption 1. The conditional expectation of demand follows the law:

$$\mathbb{E}[D_t \mid \mathcal{F}_{t-1}] = \lambda(N_{t-1}, \alpha) \cdot x(\pi_t) \tag{1}$$

for some functions λ and x. $\triangleleft$

The term $\lambda(N_{t-1}, \alpha)$ captures how the remaining inventory N_{t-1} and the cumulative sales $\alpha - N_{t-1}$ affect the expected demand in the next period, and so we call λ the *sales and inventory sensitivity* (SIS) function. We call $x(\pi_t)$ the *price sensitivity* function since it represents the effect of price on the expected demand. We assume that the seller knows the functions $\lambda(\cdot, \cdot)$ and $x(\cdot)$.

Assumption 1 states that the expected demand is of a multiplicative form which separates the effect of the current period price from the effect of past sales and inventory. Many papers use multiplicative demand functions; for instance, Smith and Agrawal (2017); Bass et al. (1994); Krishnan et al. (1999). See the review paper Urban (2005) for additional discussion. The assumption that the mean demand can depend on cumulative sales and available inventory enables us to capture situations where demand is driven by network effects (e.g., the word-of-mouth effect) or inventory availability (e.g., the scarcity effect). Table 2 summarizes the notation of our framework, working from (1) as a primitive.

Assumption 2. The SIS and price-sensitivity functions have the following properties:

(i) $x : [0, \bar{\pi}] \mapsto (0, 1]$ where $\bar{\pi}$ is an upper bound on the price. The set of feasible prices is $[0, \bar{\pi}]$,

(ii) x is continuously differentiable with derivative x' on $[0, \bar{\pi}]$ and strictly decreasing on $[0, \bar{\pi}]$ (i.e., $x'(\pi) < 0$ for all $\pi \in [0, \bar{\pi}]$). This implies that the inverse $x^{-1} : (0, 1] \mapsto [0, \bar{\pi}]$ exists and is a decreasing function,

(iii) $\pi + \frac{x(\pi)}{x'(\pi)}$, is increasing in π.

(iv) $\rho(\pi) \triangleq \pi x(\pi)$ is continuously differentiable in π and $\rho''(\pi)$ exists for all $\pi \in [0, \bar{\pi}]$,

(v) $\lambda : [0, \infty)^2 \mapsto [0, \overline{\lambda}]$ for some $\overline{\lambda} > 0$, and $\lambda(n, \alpha) > 0$ for any $0 < n \leq \alpha$,

(vi) λ is jointly concave and continuously differentiable in both of its arguments, and

(vii) $\pi_\ell(n) \triangleq x^{-1}(n/\lambda(n, \alpha))$ is differentiable in n for $n \in [0, \infty)$. $\triangleleft$

Assumption 2(i)-(iv) are standard properties of a price sensitivity function in the revenue management literature. The condition in Assumption 2(i) that $x(\pi) \leq 1$ is without loss of generality since, if it does not hold, we can simply scale the λ function correspondingly. Since $x \in [0, 1]$, then $x(\pi_t)$ essentially scales down the maximum expected demand $\lambda(N_{t-1}, \alpha)$ according to the price π_t. Assumption 2(iii) is common in the inventory and revenue management literatures, as it facilitates establishing the concavity of value functions (for a discussion, see Ziya et al. 2004; Lariviere 2006). Here, $\pi + \frac{x(\pi)}{x'(\pi)}$ is associated with the virtual value function in the mechanism design literature. If $F(\cdot)$ is the cumulative distribution function of customer valuations and $f(\cdot)$ is the associated density function, then $x(\pi)$ acts similarly to $1 - F(\pi)$ in scaling demand. Hence, $\pi + \frac{x(\pi)}{x'(\pi)} = \pi - \frac{1-F(\pi)}{f(\pi)}$ where the right-hand-side of this equation is the virtual value function, which is the virtual value of the marginal demand resulting from a marginal price change to π. Assumption 2(iv) implies that the effective revenue rate ρ is a strictly concave function and so has a unique maximizer.

Assumption 2(v)-(vi) are not restrictive since they admit a wide range of applications. Many existing demand models satisfy the concavity assumption. Some examples include the sales-dependent demand model (Bass 1969; Bass et al. 1994) and its variations (Shen et al. 2011, 2014), as well as the inventory-dependent demand models used by Datta and Pal (1990); Gerchak and Wang (1994); Urban and Baker (1997); Wang and Gerchak (2001); Sapra et al. (2010); Yang and Zhang (2014); Smith and Agrawal (2017). Moreover, all the existing papers assuming demand follows a homogeneous Poisson process satisfy our condition, e.g., Jasin and Kumar (2012, 2013); Jasin (2014); Gallego and Van Ryzin (1994, 1997); Bumpensanti and Wang (2020); Lei et al. (2021), etc.

Finally, Assumption 2(vii) is an assumption on both λ and x. It states that the lowest price $\pi_\ell(n)$ that can be charged without stocking out a supply of n in expectation is differentiable in n. It implies x and λ are smooth and guarantees the tractability of the pricing problem.

Under Assumption 2, we prove several properties that will be useful to establish the concavity of the deterministic problem and the uniqueness of its solution. The proof of the following lemma can be found in Appendix B.

Lemma 1. Define $r(y) := x^{-1}(y)y$. Under Assumption 2, the following hold:

(i) $r(y)$ is continuously differentiable, strictly concave and r'' exists for all $y \in [0, 1]$,

(ii) there exists a unique optimal solution $\bar{y}$ to the optimization problem $\max_{y \in [0,1]} r(y)$, and

(iii) $y_h(n) \triangleq n/\lambda(n,\alpha)$ is differentiable in n for $n \in [0,\alpha]$. ($y_h(n)$ is the highest intensity not causing lost sales in expectation.)

While technical in nature, Assumptions 1 and 2 are satisfied by a variety of demand models studied in the literature. Some of the demand models we explore have foundations in a consideration of willingness to pay. For example, Song and Chintagunta (2003), start with utility primitives to establish a microfoundation for the Bass model. We illustrate this in the following examples.

Example 1 (Sales-dependent demand)**.** The generalized Bass model (Bass et al. 1994; Krishnan et al. 1999) describes demand that is influenced by customers who have previously bought the product. Given a population of size k, the expected demand under this model is $\mathbb{E}[D_t \mid \mathcal{F}_{t-1}] = \lambda\left(N_{t-1},\alpha\right)x_t$, where

$$\lambda(N_{t-1},\alpha) = (k-\alpha+N_{t-1})\left(p + q\cdot\frac{\alpha - N_{t-1}}{k}\right), \tag{2}$$

and x_t captures the effect of advertising or price on the average demand. If $x_t = x(\pi_t)$ is a time-stationary function of price, then it is a price sensitivity function of the form we study in this paper. Existing literature usually assumes the price sensitivity function x takes the form of an exponential (Shen et al. 2014) or linear (Raman and Chatterjee 1995) function. In both these cases, x is consistent with Assumption 2. Note that λ in (2) also satisfies Assumption 2. ◁

Example 2 (Scarcity effect on demand)**.** Yang and Zhang (2014) and Sapra et al. (2010) model the scarcity effect in an additive demand model. Note that the assumptions used in their paper satisfy all of Assumption 2, but their demand format is in additive form, thus violating Assumption 1. However, the multiplicative version of Yang and Zhang (2014) fits our framework and assumptions. To see this, the expected demand (written in our notation) is $\mathbb{E}[D_t \mid \mathcal{F}_{t-1}] = \lambda(N_{t-1})x(\pi)$, where $\lambda(N_{t-1})$ is twice differentiable and concave decreasing in the remaining inventory N_{t-1}. The scarcity effect is captured since λ is decreasing in N_{t-1}. ◁

Example 3 (Display effect on demand)**.** Smith and Agrawal (2017) model inventory display effects through the expected demand function $\mathbb{E}[D_t \mid \mathcal{F}_{t-1}] = \lambda(N_{t-1})x(\pi)$.[2] The display effect is captured by the fact that λ is an increasing function of N_{t-1}. A canonical case that leads to several analytic results in Smith and Agrawal (2017) can be adapted to our framework with minor modifications as follows:

$$\lambda(N_{t-1}) = k(N_{t-1}/(kN_r))^{\beta} \tag{3a}$$
$$x(\pi) = e^{-\gamma\pi/c_e}, \tag{3b}$$

[2] Smith and Agrawal (2017) consider a multi-location inventory model where inventory is sold to customers in multiple locations and the seller must decide how to allocate a fixed inventory between locations. Our model is for a single location, so we adapt the single-location development (in Section 1) of Smith and Agrawal (2017). Focusing on Smith and Agrawal (2017) was largely an arbitrary choice, any number of display effect demand models could have been set into our framework (for example, Kopalle et al. 1999; Wang and Gerchak 2001).

where k is a market size, N_r and c_e are reference values, and $0 < \beta < 1$ and $\gamma > 0$. Note that λ is concave, reflecting a diminishing marginal rate of return. We can easily verify that these choices for λ and x satisfy Assumption 2. $\triangleleft$

While Assumptions 1 and 2 outline the conditions of the conditional expectation of demand, we also make the following assumption on the variance.

Assumption 3. There exists a constant $\sigma \geq 0$ such that the conditional variance of D_t for every period t does not exceed $\sigma\mathbb{E}(D_t \mid \mathcal{F}_{t-1})$.

Assumption 3 implies that, relative to the mean, the variance of demand does not become too large. Assumption 3 plays an important role in establishing the asymptotic optimality of our certainty-equivalent policies. In our asymptotic regime, the conditional expected demand scales linearly in m, and so Assumption 3 implies that the conditional standard deviation scales in order $\sqrt{m}$. Assumption 3 is reminiscent of an implicit assumption that $\text{Var}(D_t)$ scales in order m in the vast majority of existing papers on dynamic pricing papers where (independent) demand follows a Poisson or Bernoulli process (for example, Gallego and Van Ryzin 1994, 1997; Jasin and Kumar 2013; Jasin 2014). Since our setting is for state-dependent demand, Assumption 3 adapts and formalizes this to a bound on the *conditional* variance, which our analysis shows to be important for rigorously proving the revenue loss. Note the distinction between unconditional and conditional variance since $\text{Var}(D_t) = \mathbf{E}(\text{Var}(D_t \mid D_{t-1})) + \text{Var}(\mathbf{E}(D_t \mid D_{t-1}))$.

As we show in the following example, many variations of demand models where underlying randomness is governed by normal distributions, Poisson processes, and Markov chains satisfy Assumption 3. If $\sigma = 0$ then demand is deterministic, which is a special case.

Example 4. The following are a few distributions that satisfy Assumption 3:

(a) $D_t = \lambda(N_{t-1}, \alpha)x(\pi_t) + \epsilon_t$, where ϵ_t is a random component that has a normal distribution with zero mean and variance σ,

(b) D_t is a non-homogeneous Poisson process with mean $\lambda(N_{t-1}, \alpha)x(\pi_t)$,

(c) D_t is a Poisson process with constant arrival rate λ, and

(d) D_t is an aggregation of a continuous-time Markov chain with transition rate $\lambda(N_{t-1}, \alpha)x(\pi_t)$.

A final assumption (Assumption 4) on the SIS function λ relates to how it scales in an asymptotic regime. We will only present this assumption in Section 4.2 when we first introduce our asymptotic regime.

2.2 The dynamic pricing problem

Starting with initial inventory α, the seller chooses a price for each price review period based on the state. (We call this a pricing policy.) By Assumption 1, the remaining inventory N_{t-1} is sufficient to describe the state of the system at time t. Formally, a *pricing policy* $\boldsymbol{\pi} : [0, \infty) \times \{1, \ldots, T\} \mapsto \mathbb{R}_+$ (where $\mathbb{R}_+$ is the set of nonnegative real numbers) determines the price $\pi_t = \boldsymbol{\pi}(N_{t-1}, t)$ to charge at review period t given state N_{t-1}. The seller chooses an $\mathcal{F}_t$-adapted pricing policy $\boldsymbol{\pi}$ to influence the demand during the selling horizon. The expected total revenue of a pricing policy $\boldsymbol{\pi}$ is

$$V^{\boldsymbol{\pi}}(T) = \mathbb{E}\left[\sum_{t=1}^{T} \boldsymbol{\pi}(N_{t-1}, t)\left(D_t - [D_t - N_{t-1}]^+\right)\right]. \tag{4}$$

Due to periodic pricing, demand D_t can exceed the inventory N_{t-1}. Hence, a demand censoring term is included in the objective function as the total sales cannot exceed the remaining inventory. It means that, at each period, the revenue is earned only on actual sales $\min(N_{t-1}, D_t) = D_t - [D_t - N_{t-1}]^+$. In the next period, the seller will start with the remaining inventory $N_t = [N_{t-1} - D_t]^+$ for all $t \geq 1$, where $N_0 = \alpha$. The expectation in (4) is taken with respect to a stochastic demand process that is consistent with Assumptions 1 to 3. Since we examine how the number of price review periods (T) affects the algorithm and resulting profits, we do not suppress T in our notation.

Using the properties of the price sensitivity function x, we can recast the seller's decision problem. Assumption 2(ii) allows us to introduce a new variable $y_t = x(\pi_t)$ called the induced demand intensity at price π_t (or simply *intensity*) at review period t. Its inverse $\pi_t = x^{-1}(y_t)$ is uniquely determined by the intensity y_t. Thus, every pricing policy $\boldsymbol{\pi}$ has an equivalent demand intensity policy $\mathbf{y} : [0, \infty) \times \{1, \ldots, T\} \mapsto (0, 1]$. Note that for any intensity policy $\mathbf{y}$, we have

$$\mathbb{E}\left[D_t \mid \mathcal{F}_{t-1}\right] = \lambda(N_{t-1}, \alpha)\mathbf{y}(N_{t-1}, t), \quad \text{for all } t = 1, \ldots, T. \text{ (Assumption 1)} \tag{5}$$

As in the existing literature (e.g. Gallego and Van Ryzin 1994), intensity control problems are easier to analyze than pricing problems, and so we recast the problem as one where the seller is choosing an intensity policy. The expected revenue of an intensity policy $\mathbf{y}$ is

$$V^{\mathbf{y}}(T) \triangleq \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}\left(N_{t-1}, t\right)\right)\left(D_t - [D_t - N_{t-1}]^+\right)\right]. \tag{6}$$

To complete the description of the seller's problem, we now define the set of candidate (feasible) intensity policies. We let $\mathbf{Y} \triangleq \{\mathbf{y} : [0, \infty) \times \{1, \ldots, T\} \to (0, 1] \mid \mathcal{F}_t\text{-adapted}\}$ denote the set of all feasible intensity policies. The seller's problem is to choose a feasible intensity policy (and thus pricing policy) to maximize the expected revenue:

$$V^*(T) \triangleq \max_{\mathbf{y} \in \mathbf{Y}} V^{\mathbf{y}}(T). \tag{$\mathbf{P}$}$$

We denote the optimal value of this optimization problem (**P**) by $V^*(T)$.

3 Certainty-equivalent policies

Solving the stochastic pricing problem (**P**) requires knowing and working with the full demand distribution of all states. In this section, we consider certainty-equivalent (CE) policies that rely on solving a deterministic counterpart of the stochastic pricing problem (**P**).

3.1 A deterministic optimization model

We first introduce a deterministic optimization model referred to as problem ($\mathbf{D}^\dagger$):

Problem $\mathbf{D}^\dagger$

$$V^{\mathsf{D}^\dagger}(T;u,\alpha) \triangleq \max_{\substack{n\in\mathbb{R}^{T+1}\\ y\in\mathbb{R}^T}} \sum_{t=1}^{T} x^{-1}(y_t)\min\left(\lambda(n_{t-1},\alpha)y_t, n_{t-1}\right) \tag{$\mathbf{D}^\dagger$a}$$

$$\text{s.t.}\quad n_t = [n_{t-1} - \lambda(n_{t-1},\alpha)y_t]^+ \quad \text{for all } t = 1,\ldots,T \tag{$\mathbf{D}^\dagger$b}$$

$$n_0 = u \tag{$\mathbf{D}^\dagger$c}$$

$$y_t \in (0,1] \text{ for all } t = 1,\ldots,T. \tag{$\mathbf{D}^\dagger$d}$$

Note u and α are parameters of ($\mathbf{D}^\dagger$), and we assume that $0 \le u \le \alpha$. Here, u and α can both be interpreted as inventory levels. Whenever $u = \alpha$, we can check that ($\mathbf{D}^\dagger$) is a deterministic relaxation of ($\mathbf{P}$), where we replace all random variables D_t with their expectations $\lambda(n_{t-1},\alpha)y_t$. While ($\mathbf{P}$) finds an intensity policy function $\mathbf{y} : [0,\infty) \times \{1,\ldots,T\} \to (0,1]$, model ($\mathbf{D}^\dagger$) determines a vector of intensities $y = (y_1, y_2, \ldots, y_T)$. Here, $n = (n_1, n_2, \ldots, n_T)$ is the vector of remaining inventories under the deterministic demand model. Note that problem ($\mathbf{D}^\dagger$) allows $u < \alpha$ since we will later introduce a closed-loop CE policy that re-solves ($\mathbf{D}^\dagger$) in each period with the updated remaining inventory level u (with $u < \alpha$).

The objective function ($\mathbf{D}^\dagger$a) contains censored terms; hence it is non-differentiable. Further, ($\mathbf{D}^\dagger$b) is a non-convex constraint. The lack of differentiability and convexity makes problem ($\mathbf{D}^\dagger$) difficult to solve. However, we will overcome this difficulty by showing that ($\mathbf{D}^\dagger$) is equivalent to the following deterministic problem ($\mathbf{D}$) without censoring terms:

Problem D

$$V^{\mathsf{D}}(T;u,\alpha) \triangleq \max_{\substack{\mathbf{n}\in\mathbb{R}^{T+1}\\ \mathbf{y}\in\mathbb{R}^T}} \sum_{t=1}^{T} x^{-1}(y_t)\lambda(n_{t-1},\alpha)y_t \tag{$\mathbf{D}$a}$$

$$\text{s.t.}\quad \sum_{t=1}^{T} \lambda(n_{t-1},\alpha)y_t \le u \tag{$\mathbf{D}$b}$$

$$n_t = n_{t-1} - \lambda(n_{t-1},\alpha)y_t \quad \text{for all } t = 1,\ldots,T \tag{$\mathbf{D}$c}$$

$$n_0 = u \tag{$\mathbf{D}$d}$$

$$n_t \ge 0 \text{ for all } t = 0,1,\ldots,T \tag{$\mathbf{D}$e}$$

$$y_t \in (0,1] \text{ for all } t = 1,\ldots,T. \tag{$\mathbf{D}$f}$$

Note that ($\mathbf{D}$) has an additional constraint ($\mathbf{D}$b) which excludes any solutions (n,y) where the total demand exceeds inventory u. The equivalence between ($\mathbf{D}$) and ($\mathbf{D}^\dagger$) is established in the following result:

Proposition 1. For any T and $0 \leq u \leq \alpha$, the following holds:

$$V^{\mathsf{D}}(T; u, \alpha) = V^{\mathsf{D}^{\dagger}}(T; u, \alpha).$$

Moreover, finding an optimal solution to (**D**) suffices to solve (**D**†).

Proposition 1 implies that it suffices to solve problem (**D**) as the deterministic relaxation of the stochastic problem (**P**). Notice that problem (**D**) is an easier problem to solve since the objective function (**D**a) of problem (**D**) does not have demand censoring terms causing non-differentiability. We will refer to the optimal value of (**D**) when $u = \alpha$ simply as $V^{\mathsf{D}}(T)$ to be consistent with the fact that the optimal value of (**P**) is $V^*(T)$.

At first glance, the deterministic problem in (**D**) is not necessarily a convex optimization problem since the objective function is not concave and the constraints are nonlinear in the decision variables (n, y). This contrasts with the setting of Gallego and Van Ryzin (1994) where $\lambda(n_t, \alpha)$ is a constant for all n_t, resulting in a concave objective function and linear constraints. However, we can reformulate (**D**) into an equivalent convex optimization problem with decision variables $d_1, \ldots, d_T$ through a transformation:

$$d_1 = \lambda(u, \alpha) y_1, \tag{9a}$$

$$d_2 = \lambda(u - d_1, \alpha) y_2, \tag{9b}$$

$$d_3 = \lambda(u - d_1 - d_2, \alpha) y_3, \tag{9c}$$

$$\vdots$$

$$d_T = \lambda(u - d_1 - d_2 - \ldots - d_{T-1},\ \alpha) y_T. \tag{9d}$$

Here, d_t can be interpreted as the deterministic demand in period t, which depends on the amount of inventory remaining after previous periods, $u - d_1 - d_2 - \ldots - d_{t-1}$. This allows us to reformulate (**D**) into the following optimization problem, which we refer to as (**D**$'$):

Problem D$'$

$$V^{\mathsf{D}}(T; u, \alpha) = \max_{\mathbf{d} \in \mathbb{R}^T} \sum_{t=1}^{T} x^{-1}\left(\frac{d_t}{\lambda(u - d_1 - d_2 - \ldots - d_{t-1}, \alpha)}\right) \cdot d_t \tag{D$'$a}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} d_t \leq u \tag{D$'$b}$$

$$d_t \in [0, \lambda(u - d_1 - d_2 - \ldots - d_{t-1}, \alpha)] \text{ for all } t = 1, \ldots, T. \tag{D$'$c}$$

Proposition 2. The following results hold:

(i) The objective function (**D**$'$a) is jointly concave in d, and the set of all solutions satisfying constraints (**D**$'$b)–(**D**$'$c) is a convex set.

(ii) The value function $V^{\mathsf{D}}(T; u, \alpha)$ is strictly jointly concave in (u, α) for every fixed T.

Observe that (**D**$'$) is always feasible since the solution d where $d_t = 0$ for all t is feasible.

(Note that by Assumption 2(ii), an intensity 0 is in the domain of x^{-1}.) Moreover, from our continuity assumptions on x and λ, the feasible region of ($\mathbf{D'}$) is nonempty and compact, and the objective function ($\mathbf{D'}$a) is continuous, so at least one optimal solution exists (by Weierstrass's Theorem). In fact, ($\mathbf{D'}$) has a unique optimal solution, which we establish in Proposition 3.

Proposition 3 (Uniqueness)**.** For any (u, α) and T with $0 \leq u \leq \alpha$, problem ($\mathbf{D'}$) has a unique optimal solution $\mathbf{d}^{\mathsf{D}} = (d_1^{\mathsf{D}}, d_2^{\mathsf{D}}, \ldots, d_T^{\mathsf{D}})$.

Proposition 2 implies that ($\mathbf{D'}$) can be solved efficiently by any standard convex optimization algorithm. In the numerical studies we conduct, we program an interior point method to determine the optimal solution. We can also establish structure on optimal solutions to ($\mathbf{D}$).

Proposition 4 (Positive intensity is optimal)**.** The optimal solution $(\mathbf{n}^{\mathsf{D}}, \mathbf{y}^{\mathsf{D}})$ of the deterministic problem ($\mathbf{D}$) has the following properties:

(i) the remaining inventory $\mathbf{n}^{\mathsf{D}}$ is a strictly decreasing sequence, and

(ii) the optimal intensities $\mathbf{y}^{\mathsf{D}}$ are strictly positive.

The fact that remaining inventory is strictly decreasing is a sanity check since this must occur in practice (we do not model returns of the good). It is worth noting, however, that optimal intensities need not be monotone due to the potential presence of both scarcity and word-of-mouth effects of demand.

3.2 Two certainty-equivalent policies

We next introduce two certainty-equivalent (CE) policies that can be implemented by utilizing the solution of the deterministic model ($\mathbf{D}$) to set the intensity in each period. The fact that the reformulated problem ($\mathbf{D'}$) is well-behaved (Proposition 2) implies that the CE policies can be computed efficiently.

We first describe an open-loop certainty-equivalent policy (CE-OL). "Open-loop" refers to the fact that we only solve the deterministic relaxation ($\mathbf{D'}$) once (with $u = \alpha$) at the beginning of the selling horizon (time 0). After finding the optimal vector $\mathbf{y}^{\mathsf{D}}$ to ($\mathbf{D}$), the open-loop certainty-equivalent intensity policy $\mathbf{y}^{\mathsf{OL}}$ is determined by setting $\mathbf{y}^{\mathsf{OL}}(N_{t-1}, t) = y_t^{\mathsf{D}}$ for all inventory levels $N_{t-1} \in [0, \alpha]$ and $t = 1, \ldots, T$. Algorithm 1 below describes the CE-OL policy.

Algorithm 1 Intensity (price) sequence when applying policy $\mathbf{y}^{\mathsf{OL}}$.

```
procedure Open-loop Certainty Equivalent Pricing(α, T)
    d^D ← optimal solution of (D') with u = α
    for t ← 1 to T do
        y_t^D ← d_t^D / λ(α − d_1^D − d_2^D − ... − d_{t−1}^D, α)
        set intensity y_t^D by offering price x^{−1}(y_t^D)        ▷ set current intensity (price)
```

By contrast, a closed-loop certainty-equivalent policy (CE-CL) re-optimizes the deterministic problem for the remaining horizon given the current state in each period and determines the price to set in each period. We denote this policy $\mathbf{y}^{\mathsf{CL}}$.

At the start of the selling horizon when the initial inventory is $N_0 = \alpha$, CE-CL chooses the same price as CE-OL by solving ($\mathbf{D}'$) with $u = \alpha$ and setting $\mathbf{y}^{\mathsf{CL}}(N_0, t = 1) = y_1^{\mathsf{D}}$. However, for the subsequent pricing periods, the two policies diverge since CE-CL determines the next price from re-optimizing ($\mathbf{D}'$) with *updated* information about the remaining inventory. In particular, suppose that at the beginning of period t, the remaining inventory is N_{t-1}. Then CE-CL will solve ($\mathbf{D}$) with $u = N_{t-1}$ and with $T - t + 1$ periods, resulting in an optimal deterministic intensity vector $\mathbf{y}^{\mathsf{D}} = (y_1^{\mathsf{D}}, y_2^{\mathsf{D}}, \ldots, y_{T-t+1}^{\mathsf{D}})$. Note that the length of this vector is $T - t + 1$, which is the number of remaining review periods. CE-CL will set intensity $\mathbf{y}^{\mathsf{CL}}(N_{t-1}, t) = y_1^{\mathsf{D}}$. Algorithm 2 below is a description of the CE-CL intensity policy.

Algorithm 2 Intensity (price) sequence when applying policy $\mathbf{y}^{\mathsf{CL}}$.

procedure Closed-Loop Certainty Equivalent Pricing(α, T)
  $N_0 \leftarrow \alpha$ ▷ initialize inventory
  for $t \leftarrow 1$ **to** T **do**
    $d^{\mathsf{D}} \leftarrow$ optimal solution of ($\mathbf{D}'$) with $u = N_{t-1}$ and $T - t + 1$ periods
    $y_1^{\mathsf{D}} \leftarrow d_1^{\mathsf{D}} / \lambda(N_{t-1}, \alpha)$
    set intensity y_1^{D} by offering price $x^{-1}(y_1^{\mathsf{D}})$ ▷ set current intensity (price)
    observe sales $\min\{D_t, N_{t-1}\}$ by the end of period t
    $N_t \leftarrow N_{t-1} - \min\{D_t, N_{t-1}\}$ ▷ update available inventory

Although the CE-CL policy requires re-solving ($\mathbf{D}'$) in every period, solving each instance of ($\mathbf{D}'$) does not require much effort because problem ($\mathbf{D}'$) is a convex optimization problem. In our numerical experiments on a Mac Studio with an Apple M1 Ultra chip, it takes less than 3 seconds to solve ($\mathbf{D}'$) with $T = 400$ using a basic interior-point algorithm coded in Python. The running times are reported below in Table 3. Note that in ($\mathbf{D}'$), the number of variables is T and the number of constraints is $T + 1$. The CE-CL policy does not require solving ($\mathbf{D}'$) in all possible states (u, T). Specifically, the CE-CL policy can be implemented by solving ($\mathbf{D}'$) on the fly at the start of each period with the current state.

T	20	50	100	200	400
Time (seconds)	0.054	0.12	0.193	0.859	2.835

Table 3: Running times

Remark 1. It is important to note that, as mentioned in Section 2.1, demand models following a homogeneous Poisson process, of course, satisfy our demand assumptions. Consequently, our analysis includes the demand and inventory scaling studied by Gallego and Van Ryzin (1997) for a fixed price policy. A key difference in our setting is that, unlike in Gallego and Van Ryzin (1997), where the analysis is simplified due to the absence of stockouts—since the inventory is depleted by one unit at the time and the seller can charge the choke price at the exact moment inventory is depleted—we must account for censored demand. This distinction will be further elaborated in Section 4.1.

4 Asymptotic analysis of certainty-equivalent policies

Our goal in this section is to analyze the performance of the two policies proposed in Section 3.2. The main result is that the CE policies are asymptotically optimal. Specifically, in the regime where the initial inventory and the expected demand both scale by a factor m, we will show that the relative revenue loss of the CE policies compared to the true (unknown) optimal revenue converges to zero with the rate $\mathcal{O}(1/\sqrt{m})$. Note that the number of price-change opportunities T remains fixed while the market size m scales.

We prove the convergence rate in two steps. The first step is to show that the optimal deterministic revenue $V^{\mathsf{D}}(T)$ is an upper bound to the (unknown) optimal stochastic revenue $V^*(T)$, where $V^{\mathsf{D}}(T)$ is the optimal value of (**D**) when $u = \alpha$, and $V^*(T)$ is the optimal value of (**P**). The second step is to establish a rate of convergence for the CE policy's expected revenue to its upper-bound $V^{\mathsf{D}}(T)$ in the asymptotic regime of increasing inventory and expected demand. Due to the non-convexity of (**P**), and the fact its formulation allows censored demand, we cannot directly use the standard techniques (e.g., Jensen's inequality, Scarf's bound) that have been used to prove these bounds for the independent demand case. We also show that our analysis is tight by deriving lower bounds on the revenue loss of the CE policies, and showing that these lower bounds match our upper bounds.

Our tight analysis reveals an interesting insight: in a setting of state-dependent demand and with limited price change opportunities, CE policies are asymptotically optimal. Moreover, re-optimization does not improve the CE revenue loss's order of convergence since both closed-loop and open-loop CE policies achieve the same order of convergence.

4.1 Upper bound on $V^*(T)$

To understand the challenge in proving that $V^{\mathsf{D}}(T)$ is an upper bound for $V^*(T)$ in our setting, consider the (standard) situation (e.g., Gallego and Van Ryzin (1994)) where the demand rate is a constant λ (independent of the state), price can be changed at any time, and a choke price exists. Since prices can always be changed, at the moment that inventory stocks out, any pricing policy can set the choke price and turn off demand. Therefore, without loss of generality, we can assume that the total demand does not exceed the initial inventory α, so $\int_0^T dD_t \leq \alpha$. We denote by $V^\lambda(T)$ the optimal expected revenue of this simplified problem. Let $\mathbf{y}^\lambda = (y_t^\lambda)$ be the optimal intensity policy. Following the proof technique of Lemma 1 in Gallego and Van Ryzin 1994, for any $\mu \geq 0$

$$\begin{aligned} V^\lambda(T) = \mathbb{E}\left(\int_0^T x^{-1}(y_t^\lambda)\mathrm{d}D_t\right) &\leq \mathbb{E}\left(\int_0^T x^{-1}(y_t^\lambda)\mathrm{d}D_t + \mu\left(\alpha - \int_0^T \lambda y_t^\lambda \mathrm{d}t\right)\right) \\ &\leq \max_{\{y_t: t\in[0,T]\}} \left(\int_0^T x^{-1}(y_t)\lambda y_t \mathrm{d}t + \mu\left(\alpha - \int_0^T \lambda y_t \mathrm{d}t\right)\right) \qquad (11) \\ &\leq \int_0^T \max_{y_t}\left(x^{-1}(y_t)\lambda y_t\mathrm{d}t + \mu\left(\alpha - \int_0^T \lambda y_t\mathrm{d}t\right)\right). \qquad (12) \end{aligned}$$

The first inequality is from Lagrangian relaxation since we know that the expected demand cannot exceed α. The second and last inequality is from maximizing pointwise for each t and by Jensen's inequality. Note that the right-hand side of (12) is the Lagrangian relaxation of the deterministic model. The deterministic counterpart is a convex optimization problem (since $x^{-1}(y_t)y_t$ is concave in y_t), so strong duality holds and the right-hand side is equal to $V^{\mathsf{D}}(T)$ when taking the infimum over $\mu \geq 0$.

This analysis does not apply to our setting. The first issue is the demand censoring term in (6). This means that the deterministic relaxation ($\mathbf{D}^{\dagger}$) is a non-convex optimization problem and strong duality does not necessarily hold. Even though Proposition 1 shows the equivalence of ($\mathbf{D}^{\dagger}$) to the model ($\mathbf{D}$) without censoring, the constraint ($\mathbf{D}$b) is still non-convex. Going to ($\mathbf{D'}$), which is convex in its decision variable $\mathbf{d}$, goes too far because we need to keep a similar formulation structure involving the policy $\mathbf{y}$. Thus, strong duality is still not guaranteed even in a model without the demand censoring terms.

A second issue comes from the fact that demand is state-dependent. As a result, the point-wise maximum in (12) cannot be taken in our setting since the expected demand in period t depends on the remaining inventory N_{t-1}, which in turn depends on previous intensities $y_1, \ldots, y_{t-1}$.

Our proof overcomes both issues by establishing the bound, not directly on ($\mathbf{D}$) and ($\mathbf{P}$), but through mathematical induction on their dynamic programming (DP) counterparts. Specifically, the DP counterpart of ($\mathbf{D}$) for any $u \in [0, \alpha]$ is:

$$\begin{aligned} R^{\mathsf{D}}(u,T) \quad &\triangleq \quad \max_{y\in[0,1]} \quad x^{-1}(y)\lambda(u,\alpha)y + R^{\mathsf{D}}\left(u - \lambda(u,\alpha)y,\ T-1\right) \\ &\qquad\quad \text{s.t.} \quad\ \lambda(u,\alpha)y \leq u, \end{aligned} \tag{13}$$

where the base case is $R^{\mathsf{D}}(u, 0) = 0$ for all $u \in [0, \alpha]$. The value $R^{\mathsf{D}}(u, T)$ can be thought of as the deterministic revenue-to-go if the remaining inventory is u and there are T price change opportunities remaining. Hence, we have $V^{\mathsf{D}}(T) = R^{\mathsf{D}}(\alpha, T)$.

Similarly, for any $u \in [0, \alpha]$, the stochastic optimization problem ($\mathbf{P}$) has a dynamic programming counterpart:

$$R^{*}(u,T) \quad \triangleq \quad \max_{y\in[0,1]} \quad \mathbb{E}_{y,u}\left[x^{-1}(y)\left(D - [D-u]^{+}\right) + R^{*}([u-D]^{+},\ T-1)\right]. \tag{14}$$

Here, $\mathbb{E}_{y,u}$ is the expectation with respect to the distribution of per-period demand D when the remaining inventory at the start of the period is u, and y is the current period intensity. Recall that y and u affect the distribution of D, including but not limited to its conditional mean $\lambda(u, \alpha)y$. The base case is $R^{*}(u, 0) = 0$ for all $u \in [0, \alpha]$. Note that $R^{*}(u, T)$ can be thought of as the optimal expected revenue-to-go if the remaining inventory is u and there are T price change opportunities remaining. Hence, $V^{*}(T) = R^{*}(\alpha, T)$.

Our focus on the DP formulations overcomes the two issues we identified at the outset of this subsection. The first issue (potential lack of strong duality) is resolved because if (u, T) is given, the constraint in $\lambda(u, \alpha)y \leq u$ in (13) is linear in y. Using mathematical induction, we can also establish that the objective of (13) is strictly concave in y due to the concavity assumption on $\lambda(\cdot,\cdot)$. Hence, strong duality holds for the Lagrangian relaxation of (13). Strong duality is the crucial step to establishing the upper bound. The second issue (inability to take a pointwise maximum) is resolved because we take the maximum of (14) only for the revenue-to-go, and the effect of current y_t on future periods is absorbed in the term $R^*([u-D]^+, T-1)$. Combining these ideas allows us to prove the upper-bound result.

Aided by the DP formulations and this proof idea, the following result establishes that the optimal expected revenue-to-go is bounded above by the deterministic revenue-to-go.

Proposition 5 (Upper bound)**.** For any $T \geq 1$, $V^*(T) \leq V^{\mathsf{D}}(T)$. More generally, for any $0 \leq u \leq \alpha$, $R^*(u,T) \leq R^{\mathsf{D}}(u,T)$.

The complete proof can be found in Appendix C.6.

4.2 Asymptotic regime

We now consider a scaled version of the problem with $m \in \mathbb{Z}^+$ as a scaling factor. First, we scale the initial inventory to αm. At the same time, for any period $t = 1, \ldots, T$, we assume that the scaled random demand, denoted as D_t^m, has a conditional mean satisfying the following assumption:

Assumption 4. The conditional expectation of the demand D_t^m has an SIS function λ^m that scales in m such that

$$\lambda^m(N_{t-1}^m, \alpha m) = m\lambda\left(\frac{N_{t-1}^m}{m}, \alpha\right), \tag{15}$$

where λ is a function that is independent of m and that satisfies Assumption 2(v)–(vi).

Here, N_{t-1}^m denotes the inventory level at the start of period t, which is a $\mathcal{F}_{t-1}$-measurable random variable. By definition, $N_0^m = \alpha m$. Assumption 4, together with Assumption 1, implies that the conditional expectation of demand scales linearly with m. Note that Assumption 4 is only required for the proof of asymptotic optimality. Assumption 4 is not restrictive and can be easily satisfied. For example, if the demand rate is a constant λ, such as in a homogeneous Poisson process, Assumption 4 holds by simply scaling the demand rate as λm.

In the demand model of Example 1, Assumption 4 holds if the market size scales as km. Indeed, from (2), we have that

$$\begin{aligned}\lambda^m(N_{t-1}^m, \alpha m) &= \left(km - \alpha m + N_{t-1}^m\right)\left(p + q\frac{\alpha m - N_{t-1}^m}{km}\right) \\ &= m\left(k - \alpha + \frac{N_{t-1}^m}{m}\right)\left(p + q\frac{\alpha - N_{t-1}^m/m}{k}\right) = m\lambda\left(\frac{N_{t-1}^m}{m}, \alpha\right).\end{aligned}$$

In the demand model of Example 3, Assumption 4 also holds when the market size scales as km. From (3a), we have

$$\lambda^m(N_{t-1}^m) = (km)\left(\frac{N_{t-1}^m}{(km)N_r}\right)^{\beta} = m \cdot k\left(\frac{(N_{t-1}^m/m)}{kN_r}\right)^{\beta} = m\lambda(N_{t-1}^m/m).$$

Additionally, Assumption 4 holds if, for all $m \in \mathbb{Z}^+$, we have $\lambda^m = \lambda$ where λ is a homogeneous function of degree 1. The property by definition means that $\lambda(N^m_{t-1}, \alpha m) = m\lambda(N^m_{t-1}/m, \alpha)$.

The scaled version of the pricing problem ($\mathbf{P}_m$) is:

$$V^*(m,T) \triangleq \max_{\mathbf{y}\in\mathbf{Y}} V^{\mathbf{y}}(m,T), \tag{$\mathbf{P}_m$}$$

where the expected revenue $V^{\mathbf{y}}(m,T)$ of policy $\mathbf{y}$ is defined as:

$$V^{\mathbf{y}}(m,T) \triangleq \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}(N^m_{t-1},t)\right)\left(D^m_t - [D^m_t - N^m_{t-1}]^+\right)\right]. \tag{16}$$

Recall that $\mathbf{Y}$ is the set of all intensity policies $\mathbf{y}$ that are $\mathcal{F}_t$-measurable. The dynamics of the remaining inventory is $N^m_t = [N^m_{t-1} - D^m_t]^+$, where $N^m_0 = \alpha m$ is the scaled initial inventory. For any m, the distribution of D^m_t satisfies Assumptions 1 to 3.

We use ($\mathbf{D}_m$) to denote the scaled counterpart of the deterministic model ($\mathbf{D}$) where α is replaced with αm and $\lambda(n_{t-1}, \alpha)$ is replaced by $\lambda^m(n_{t-1}, \alpha m)$. Per our discussion in Section 3.1, if $u = \alpha m$, then ($\mathbf{D}_m$) is the deterministic counterpart to the scaled stochastic problem ($\mathbf{P}_m$). Let $V^{\mathsf{D}}(m,T)$ denote the optimal value of ($\mathbf{D}_m$) when we set $u = \alpha m$. Note that $V^{\mathsf{D}}(1,T) = V^{\mathsf{D}}(T)$.

An immediate consequence of Proposition 5 is that $V^*(m,T) \le V^{\mathsf{D}}(m,T)$. The implication of this is that a policy $\mathbf{y}$ is asymptotically optimal if, as m increases, the bound on its expected revenue loss, $V^{\mathsf{D}}(m,T) - V^{\mathbf{y}}(m,T)$, grows at a slower rate than the growth rate of $V^{\mathsf{D}}(m,T)$. Note that $V^{\mathsf{D}}(m,T)$ grows linearly in m. This is because (15) implies $\lambda^m(mn, m\alpha) = m\lambda(n,\alpha)$ for any $n \in [0,\alpha]$. Hence, when we set $u = \alpha m$ for ($\mathbf{D}_m$) and $u = \alpha$ for ($\mathbf{D}$), we can check that their respective optimal solutions, $(\mathbf{n}^{\mathsf{D},m}, \mathbf{y}^{\mathsf{D},m})$ and $(\mathbf{n}^{\mathsf{D}}, \mathbf{y}^{\mathsf{D}})$, have the property that $\mathbf{n}^{\mathsf{D},m} = m\mathbf{n}^{\mathsf{D}}$ and $y^{\mathsf{D},m} = \mathbf{y}^{\mathsf{D}}$. This implies that $V^{\mathsf{D}}(m,T) = mV^{\mathsf{D}}(T)$, hence the linear growth of $V^{\mathsf{D}}(m,T)$.

We will analyze the convergence rate of the expected revenue loss under our proposed policies, $\mathbf{y}^{\mathsf{OL}}$ and $\mathbf{y}^{\mathsf{CL}}$. For scaling factor m, $\mathbf{y}^{\mathsf{OL}}$ and $\mathbf{y}^{\mathsf{CL}}$ are based on solutions to the scaled model ($\mathbf{D}_m$) instead of ($\mathbf{D}$). Given m, let $V^{\mathsf{OL}}(m,T)$ and $V^{\mathsf{CL}}(m,T)$ denote the expected revenue under the CE-OL and CE-CL, respectively. Hence, the expected revenue losses under CE-OL and CE-CL are $V^{\mathsf{D}}(m,T) - V^{\mathsf{OL}}(m,T)$ and $V^{\mathsf{D}}(m,T) - V^{\mathsf{CL}}(m,T)$, respectively. In Section 4.3, we show that both expected revenue losses are lower bounded by $\Omega(\sqrt{m})$. Then, in Section 4.4 we show that both expected revenue losses are upper bounded by $\mathcal{O}(\sqrt{m})$. Hence, the CE policies are asymptotically optimal as m grows large since the relative revenue loss compared to the true (unknown) optimal policy is $\mathcal{O}(1/\sqrt{m})$.

Showing that $V^{\mathsf{D}}(m,T) - V^{\mathsf{OL}}(m,T)$ and $V^{\mathsf{D}}(m,T) - V^{\mathsf{CL}}(m,T)$ are both $\mathcal{O}(\sqrt{m})$ does not immediately follow from standard arguments in the existing literature (e.g., Gallego and Van Ryzin 1994; Jasin 2014). This is because, in our setting, the demand is a random variable that depends on the path of remaining inventory through the function λ. Therefore, the deviation of the expected revenue from $V^{\mathsf{D}}(m,T)$ does not just depend on the expected stock-out level, it also depends on deviations of the path of remaining inventory from the optimal inventory solution $(n_0^{\mathsf{D},m},\ldots,n_T^{\mathsf{D},m})$ of the deterministic counterpart $(\mathbf{D}_m)$ when $u=\alpha m$. Hence, it is crucial to establish the convergence of the remaining inventory paths to their deterministic equivalents (see Lemmas 2 and 7 below). Assumption 3 is crucial for this step since it implies that the variance does not grow too fast as the problem scales up, so the normalized demand D_t^m/m can be well approximated by its mean as m scales up. Most notably, the demand paths and inventory paths under the certainty-equivalent policies also converge to the deterministic optimal path, making the relative revenue losses of both CE policies converge to zero.

4.3 Lower bound on CE expected revenue loss

The next result establishes a fundamental lower bound of $\Omega(\sqrt{m})$ on the expected revenue loss of any feasible policy. The example comes from the proof of Proposition 2 in Besbes and Zeevi (2009).

Theorem 1. Let $\lambda(u,\alpha)=1, \lambda^m(mu,\alpha m)=m\lambda(u,\alpha)$ for any u,α, and let Assumption 2 hold with $x(\pi)=a-b\pi$ where $b\in[\underline{b},\bar{b}], \underline{b}>0$ and $a\in\max\{2\bar{b}\bar{\pi},\bar{b}\bar{\pi}+\alpha/T\}$. Per period demand is independently distributed following a Poisson distribution. For any policy $\mathbf{y}$, the following holds:

$$V^{\mathsf{D}}(m,T)-V^{\mathbf{y}}(m,T)\geq\Omega(\sqrt{m}).$$

The proof of Theorem 1 written in our notation can be found in Appendix C.7. We note that the lower bound result holds if initial inventory α and number of price changes T grow at the same rate (but the length of total selling horizon is fixed), and demand is at most one in each period.

At first glance, Theorem 1 seems to contradict a main result of Jasin (2014). Jasin (2014) shows that when demand follows a Poisson process, in a continuous time setting (at most one demand per period), the CE-CL policy has an $\mathcal{O}(\log m)$ bound on the expected revenue loss, which is better than the $\Omega(\sqrt{m})$ lower bound derived here.[3] However, the upper bound result in Jasin (2014) is shown under the condition that more inventory strictly improves the revenue (condition $\mu^D>0$ in Theorem 1 of Jasin 2014). In the instance in Theorem 1, the condition that more inventory strictly improves the revenue fails. It is this extra condition that drives the $\log(m)$. The other properties of the model (such as limited price changes and dependent demand) are not drivers. In particular, the constructed example in Theorem 1 assumes independent demand over time. Therefore, it is not the lack of independence that leads to the square-root lower bound.

With regards to the extra condition in Jasin (2014) that inventory strictly improves revenue, in our setting of state-dependent demand, more inventory could result in strictly lower revenue. An example where this could happen is when scarcity boosts sales, so higher inventory results in a lower demand rate. It would be interesting to look into the question of whether re-optimization

[3] Jasin (2014) considers the asymptotic regime that scales up the total planning horizon (and, therefore, the number of price changes T) but fixes the demand rate at $\lambda x(\pi)$ during the unit time δt. All their asymptotic results also hold if they keep the length of the selling horizon fixed while scaling up the demand rate to $m\lambda x(\pi)$.

gives $o(\sqrt{m})$ revenue loss under the condition that more inventory strictly improves revenue in our setting. We attempted this, but the approach in Jasin (2014) requires explicit characterizations of the optimal primal and dual solutions of the re-optimized deterministic problem. Such explicit characterization are not easy to find in our setting due to the nonlinear sales and inventory dependent $\lambda(\cdot,\cdot)$ function.

4.4 Upper bound on CE expected revenue loss

We next show that the expected revenue loss of the open-loop policy, $V^D(m,T) - V^{\mathsf{OL}}(m,T)$, and of the closed-loop policy, $V^D(m,T) - V^{\mathsf{CL}}(m,T)$, both grow in order $\mathcal{O}(\sqrt{m})$ as m scales. Hence our lower bound result (Theorem 1) implies that both certainty equivalent policies have an expected revenue loss that is $\Theta(\sqrt{m})$. That is, re-optimization does not improve the CE revenue loss' asymptotic order of growth.

We begin by analyzing the loss under the open-loop policy. We introduce some notation. Observe that the open-loop policy $\mathbf{y}^{\mathsf{OL}}$ is a static, but time-varying policy. Thus, we use y_t^{OL} to denote the *deterministic* period t intensity using the open-loop policy $\mathbf{y}^{\mathsf{OL}}$.[4] For a given m, let $\bar{N}^m = (\bar{N}_0^m, \ldots, \bar{N}_T^m)$ be the stochastic sequence of inventory levels under the open-loop certainty-equivalent policy $\mathbf{y}^{\mathsf{OL}}$. Note that $\bar{N}_0^m = \alpha m$.

The next lemma states that the normalized inventory $\bar{N}_t^m/m$ of the open-loop policy converges in expectation to the deterministic optimal inventory n_t^{D} solution to (**D**) when $u = \alpha$. This result is a crucial step to analyze the convergence of the expected revenue because demand is state-dependent. When λ is a constant under i.i.d demand, this convergence holds trivially. This lemma implies that the expected demand rate of the open-loop policy converges in expectation to the deterministic optimal demand rate even though the conditional expectation of demand is state-dependent. The proof of this lemma is in Appendix C.8.

Lemma 2 (Convergence of remaining inventory and SIS)**.** If $\mathbf{n}^{\mathsf{D}} = (n_1^{\mathsf{D}}, \ldots, n_T^{\mathsf{D}})$ is the solution of remaining inventories to (**D**) when $u = \alpha$, then the following hold:

$$\mathbb{E}\left|\frac{\bar{N}_t^m}{m} - n_t^{\mathsf{D}}\right| = \mathcal{O}\left(1/\sqrt{m}\right), \quad \text{for all } t = 1, \ldots, T, \tag{17}$$

$$\mathbb{E}\left|\lambda\left(\frac{\bar{N}_t^m}{m}, \alpha\right) - \lambda\left(n_t^{\mathsf{D}}, \alpha\right)\right| = \mathcal{O}\left(1/\sqrt{m}\right), \quad \text{for all } t = 1, \ldots, T. \tag{18}$$

With the help from Lemma 2, we are able to show that the difference from $V^{\mathsf{D}}(m,T)$ of the expected *uncensored* revenue of $\mathbf{y}^{\mathsf{OL}}$ is order $\mathcal{O}(\sqrt{m})$. The uncensored revenue (corresponding to the first term in (19) below) is computed, assuming all demands can be sold irrespective of the inventory level. The proof is in Appendix C.9.

Lemma 3 (Convergence of uncensored revenue)**.** The following holds:

$$\left|\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda^m\left(\bar{N}_{t-1}^m, \alpha m\right) y_t^{\mathsf{OL}}\right) - V^{\mathsf{D}}(m,T)\right| = \mathcal{O}\left(\sqrt{m}\right). \tag{19}$$

[4]Since it is open-loop, y_t^{OL} is independent of demand realization

Though the bound in Lemma 3 is for an uncensored setting, we use this result to derive the loss bound for the expected revenue in the censored setting.[5] This, combined with Proposition 5, establishes the asymptotic bound for the expected revenue loss of $\mathbf{y}^{\mathsf{OL}}$ (Theorem 2 below). Specifically, the proof of the next result (in Appendix C.10) first shows that the censored revenue $V^{\mathsf{OL}}(m,T)$ converges to the uncensored revenue as m grows large and then uses Lemma 3 to show the result.

Theorem 2 (Expected revenue loss of open-loop CE policy)**.** Under Assumptions 1 to 4, the following holds:

$$1-\frac{V^{\mathsf{OL}}(m,T)}{V^*(m,T)} \le 1-\frac{V^{\mathsf{OL}}(m,T)}{V^{\mathsf{D}}(m,T)} = \mathcal{O}\left(1/\sqrt{m}\right). \tag{20}$$

The implication of Theorem 2 is that the open-loop policy performs well if the problem scale m is large. It is important to note that the asymptotic optimality result of Theorem 2 applies for any demand distribution.

The analysis of the expected revenue loss under the closed-loop policy $\mathbf{y}^{\mathsf{CL}}$proceeds similarly to that of $\mathbf{y}^{\mathsf{OL}}$ except with one key difference. The difference is that we need to show $\mathbf{y}^{\mathsf{CL}}(n,t)$ is Lipschitz continuous in any $n \in [0, \alpha m]$. This is formalized in the following lemma.

Lemma 4 (Lipschitz continuous policy)**.** There exists a uniform constant C_y such that, for any $n, n' \ge 0$,

$$\left|\mathbf{y}^{\mathsf{CL}}(n,t) - \mathbf{y}^{\mathsf{CL}}(n',t)\right| \le C_y \left|n-n'\right|, \quad \text{for all } t = 1,\dots,T.$$

This property is important since, unlike the open-loop policy that has a static price sequence, $\mathbf{y}^{\mathsf{CL}}$ results in a stochastic price sequence that dynamically changes based on the past realizations of demand. Since $\mathbf{y}^{\mathsf{CL}}$ is a Lipschitz continuous function in n, then the difference in price at two inventory levels does not grow too fast, compared to the difference in inventory level. This is desirable since it leads to a relatively stable pricing policy against inventory dynamics.

With this key property, we can establish convergence of the inventory sequence under $\mathbf{y}^{\mathsf{CL}}$ to the deterministic inventory sequence. This is formalized in Lemma 7, which is stated and proved in Appendix C.12. This allows us to show that the *uncensored* expected revenue under $\mathbf{y}^{\mathsf{CL}}$ has a gap from $V^{\mathsf{D}}(m,T)$ that is $\mathcal{O}(\sqrt{m})$. This is formalized in Lemma 8, which is stated and proved in Appendix C.13. Note that Lemma 7 and Lemma 8 are the counterparts of Lemma 2 and Lemma 3, respectively, for the closed-loop policy.

Hence, as with the open-loop policy, the closed-loop certainty equivalent policy $\mathbf{y}^{\mathsf{CL}}$ is asymptotically optimal to the stochastic pricing problem as the problem scale m grows large. Its proof is in Appendix C.14.

Theorem 3 (Expected revenue loss of closed-loop CE policy)**.** The following holds:

$$1-\frac{V^{\mathsf{CL}}(m,T)}{V^*(m,T)} \le 1-\frac{V^{\mathsf{CL}}(m,T)}{V^{\mathsf{D}}(m,T)} = \mathcal{O}\left(1/\sqrt{m}\right). \tag{21}$$

[5] We use the Scarf bound (Scarf 1958), which establishes the expected difference between a truncated random variable and itself, to show the difference between the censored revenue and the uncensored revenue.

The asymptotic optimality of the closed-loop policy holds for any demand distribution that satisfies Assumptions 1 to 4.

4.5 Discussion of our analysis

We would like to point out two distinctive features in our problem that make our analysis of the CE policies different from earlier works in dynamic pricing literature.

The first feature is that the demand in each period is state-dependent, hence the demands across periods are dependent. Unlike the case where demands are independent (among many examples are Gallego and Van Ryzin 1994; Maglaras and Meissner 2006; Jasin and Kumar 2013), we need to introduce new mathematical machinery to prove the asymptotic optimality of the CE policies. For example, we establish the upper bound result of Proposition 5 by converting the problem to dynamic programming formulations of (**P**) and (**D**). If the demands were independent, this upper bound can be shown by Lagrangian relaxation directly on the multi-period model. Further, in this setting, the $\mathcal{O}(\sqrt{m})$ gap between the CE policy expected revenue and the deterministic upper bound can be trivially established. But when demands are state-dependent, the $\mathcal{O}(\sqrt{m})$ bound can only be established if the expected "path" of states (i.e., the inventory level) under the CE policy converges to the optimal deterministic inventory level. This is non-trivial to show when demands are state-dependent, since the cumulative sales (and the resultant inventory level) in the previous periods affect the demand and inventory of the current period.

The second feature is that the price-change opportunities are limited. Hence, the inventory may stock out during a period, resulting in a demand censoring term in the revenue function. Censored demands make the analysis non-trivial even if the demands were independent. For example, when there is no censoring, an upper bound can be established using straightforward arguments since the deterministic relaxation is a convex problem, as we discussed in Section 4.1.

Many existing works in dynamic pricing literature assume continuous price changes (combined with Poisson demand arrivals), so without loss of generality, demand is uncensored. This is because any continuous review pricing policy can simply turn off demand by setting a high price at the exact moment that inventory reaches zero. Due to the uncensored demand, the analysis in those continuous price review models is tractable. A setting somewhat resembling ours is Section 5.1 of Gallego and Van Ryzin (1994), which considers a compound Poisson process where, at each Poisson arrival time, a random demand size is observed. However, Gallego and Van Ryzin (1994) restrict their analysis to policies where the resulting total demand does not exceed inventory almost surely, so there is no demand censoring in the objective. We make no such simplifying assumption. Our analysis of asymptotic optimality needs to hold with demand censoring. We overcome this challenge in several steps of the analysis. First, we show the connection of the censored deterministic relaxation ($\mathbf{D}^{\dagger}$) to a model (**D**) where deterministic demand cannot exceed inventory. This property of the deterministic solution is used in several places of the proofs, such as in establishing the deterministic upper bound (Proposition 5) and in proving the inventory path convergence of the CE policies (Lemmas 2 and 7). Second, we bound the difference between the censored and uncensored expected revenues by bounding the expected lost sales using Scarf (1958), as can be seen in the proofs of Theorems 2 and 3.

We also want to note in our setting, the number of pricing decisions T does not scale up with the scaling factor m in the asymptotic regime. This is different from asymptotic regimes considered in the literature (e.g., Gallego and Van Ryzin (1994); Jasin (2014)), the number of decisions scales up with the scaling factor. For example, in Jasin (2014), when m becomes large, he scales the number of opportunities with re-optimization policies to correct their pricing errors. However, in our setting with the number of price change opportunities fixed at T, we are still able to show CE policies are asymptotically optimal despite much less pricing flexibility.

We should note, however, that our asymptotic results still hold if we scale up the number of pricing decisions T. This is operationalized by shrinking the time between price changes by a factor of $1/m$ while m grows. Under this scenario, the certainty-equivalent policies are also asymptotically optimal if initial inventory α and T grow at the same rate.

Finally, we emphasize the importance of the concavity assumption on $\lambda(\cdot,\cdot)$ for CE policies to achieve asymptotic optimality. First, in order to prove the fluid upper bound (Proposition 5), we need to establish strong duality of the deterministic model. Strong duality is also needed in the independent and stationary demand case (i.e., constant λ) to bound the deterministic model. However, in that case, it is trivial to establish strong duality, and so it is typically not even mentioned as a distinct step in the analysis (as in Gallego and Van Ryzin (1994); Jasin (2014)). The concavity of λ is used to establish the strong duality of the deterministic problem in our problem. If λ is a general and not concave function, the objective function of the deterministic problem is not concave, so strong duality cannot be easily shown. Proving the fluid upper bound result is critical for proving an upper bound on revenue loss. Second, the concavity of λ also guarantees the existence of a unique solution to the deterministic problem, which implies the tractability of our CE policies. The CE policies can possibly perform worse than $\sqrt{m}$ if the CE policy is not unique (see the example in the appendix Example 5). Third, the concavity of λ guarantees the Lipschitz continuity of y^{CL} used to bound the absolute difference of the expected revenue loss under CE policies. This is not needed when λ is a constant because the CE optimal policy is a fixed-price policy.

5 Further analysis

In this section, we expand on a few more ideas related to what we presented so far. First, we ask what if the seller can also select the starting inventory level. Second, we consider the question of what is the incremental benefit of adding more price change-opportunities.

5.1 Joint optimization of starting inventory and pricing

Consider an extension where the seller sets the initial inventory along with prices. At time 0, the seller decides an initial inventory $N_0 = \alpha$ by choosing $\alpha \geq 0$, and incurs a procurement cost of c per unit of inventory.

The full stochastic version of the problem would be optimize the expected profit of decision

$(\alpha, \mathbf{y})$ given by

$$Q^{\alpha,\mathbf{y}}(T) \triangleq \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}\left(N_{t-1}, t\right)\right)\left(D_t - [D_t - N_{t-1}]^+\right)\right] - c\alpha,$$

where $N_0 = \alpha$ and $N_t = [N_{t-1} - D_t]^+$ for all $t \geq 1$. Note that $Q^{\alpha,\mathbf{y}}(T) = V^{\alpha,\mathbf{y}}(T) - c\alpha$, where we write $V^{\alpha,\mathbf{y}}(T)$ instead of $V^{\mathbf{y}}(T)$ to emphasize that α is a decision variable. Hence, with full knowledge of the demand distribution, the seller's decision problem is

$$Q^*(T) \triangleq \max_{\alpha \geq 0} \max_{\mathbf{y} \in \mathbf{Y}} Q^{\alpha,\mathbf{y}}(T). \tag{$\mathbf{P}'$}$$

The only difference from Section 2.2 is that now α is a decision variable.

We now introduce a certainty-equivalent policy that only requires knowledge of the functions λ and x that specify the conditional expectation of per-period demand. Consider the following problem:

$$Q^{\mathsf{D}}(T) \triangleq \max_{\alpha \geq 0} Q^{\mathsf{D},\alpha}(T) := \max_{\alpha \geq 0} V^{\mathsf{D},\alpha}(T) - c\alpha, \tag{$\mathbf{D}'$}$$

where we write $V^{\mathsf{D},\alpha}(T)$ instead of $V^{\mathsf{D}}(T)$ to emphasize that α is a decision variable that affects the expected revenue through the inventory constraint and in scaling the demand rate through $\lambda(n, \alpha)$. Note that $Q^{\mathsf{D},\alpha}(T)$ in ($\mathbf{D}'$) is the deterministic counterpart of $\max_{\mathbf{y} \in \mathbf{Y}} Q^{\alpha,\mathbf{y}}(T)$ in ($\mathbf{P}'$).

The certainty-equivalent policy solves the deterministic counterpart ($\mathbf{D}'$) to set the initial inventory $\alpha^{\mathsf{CE}} \geq 0$. Given $\alpha = \alpha^{\mathsf{CE}}$, the policy then sets $\mathbf{y}^{\mathsf{CE}} : [0, \infty) \times \{1, \ldots, T\} \mapsto [0, 1]$ as either one of the certainty-equivalent intensity policies described in the previous sections, where $\mathsf{CE} \in \{\mathsf{OL}, \mathsf{CL}\}$. We denote the expected profit of the certainty-equivalent policy of the joint inventory and pricing problem as $Q^{\mathsf{CE}}(T)$.

Algorithm 3 gives a description of the CE policy.

Algorithm 3: Setting initial inventory and prices with the CE policy.

procedure Certainty Equivalent(T)
    $\alpha^{\mathsf{CE}} \leftarrow$ optimal solution of ($\mathbf{D}'$)
    set $N_0 = \alpha^{\mathsf{CE}}$ ▷ set initial inventory
    set prices according to the CE-policy (open-loop or closed-loop) for $(\alpha^{\mathsf{CE}}, T)$

Computing the certainty-equivalent policy for a joint inventory and pricing policy is tractable. Recall that in Proposition 2(ii), we prove that the deterministic value function $V^{\mathsf{D}}(T; u, \alpha)$ is jointly concave in (u, α) for a given T. This implies that solving for the certainty-equivalent market coverage α^{CE} can be simply done by gradient methods like the Newton algorithm.

Consider a setting where we scale by a factor m both the initial inventory and the expected demand by (15). We denote the optimal expected profit as $Q^*(m, T)$ and the expected profit of the certainty-equivalent policy is $Q^{\mathsf{CL}}(m, T)$. As in the case with the certainty-equivalent pricing policies, we show that the expected profit loss under Algorithm 3 grows sub-linearly in m –

this means that our proposed joint decision policy is asymptotically optimal. This is formally established in Theorem 4. The proof is in Appendix D.1.

Theorem 4 (Expected profit loss of CE policies)**.** The following holds:

$$1 - \frac{Q^{\mathsf{CE}}(m,T)}{Q^{*}(m,T)} = \mathcal{O}\left(1/\sqrt{m}\right). \tag{22}$$

This result shows that the CE policy guarantees a close-to-optimal expected profit when the scale of inventory and demand is large. We note that the CE policy may choose a different initial inventory from the optimal initial inventory of the mth stochastic problem (which we denote by $\alpha^*(m)$). Thus, the asymptotic optimality in Theorem 4 does not follow immediately from Theorems 2 and 3. The implication of Theorem 4 is that when m is large enough, the scaled-down initial inventory $\alpha^*(m)/m$ is close to α^{CE}.

5.2 Analyzing the benefit of additional price-change opportunities

We are interested in understanding the benefit (in terms of asymptotic regret) of having an additional price change opportunity. To this end, we quantify the benefit of price change opportuntie by comparing a fixed price policy (single price for T periods) to a policy that allows to change price T times over T periods, which we will call a T-price policy.

The optimal fixed-price policy, denoted by $\mathbf{y}^{\mathsf{FP}}$, is as follows. If the initial inventory, α, is sufficiently large, the fixed-price policy fixes a price corresponding to intensity $\bar{y}$, the unique maximizer of the revenue function, i.e., $\bar{y} \triangleq \arg\max_{y\in[0,1]} x^{-1}(y)y$. If the inventory constraint is binding, the policy instead chooses the intensity so that the expected total demand equals the initial inventory, i.e., the fixed point y^{so} of the equation (the superscript "so" stands for "stockout price"):

$$\bar{y}^{\text{so}} = \frac{\alpha}{\sum_{t=1}^{T} \lambda\left(n_{t-1}^{\bar{y}^{\text{so}}}, \alpha\right)},$$

where, for any $y \in [0,1]$, $(n_0^y, n_1^y, \ldots, n_T^y)$ is defined as the deterministic sequence with $n_0^y = \alpha$ and $n_t^y = n_{t-1}^y - \lambda(n_{t-1}^y, y)y$ for all $t \in \{1,\ldots,T\}$. Note that $\bar{y}^{\text{so}}$ can be found by fixed point iteration. Therefore, formally, given any initial inventory $\alpha \geq 0$, the optimal fixed-price policy $\mathbf{y}^{\mathsf{FP}}$ is defined for every $(n,t) \in (0,\alpha] \times \mathcal{T}$ as:

$$\mathbf{y}^{\mathsf{FP}}(n,t) = y^{\mathsf{FP}} \triangleq \begin{cases} \bar{y}, & \text{if } \alpha \geq \sum_{t=1}^{T} \lambda\left(n_{t-1}^{\bar{y}}, \alpha\right)\bar{y}, \\ \bar{y}^{\text{so}}, & \text{otherwise.} \end{cases} \tag{23}$$

To implement this, the seller will charge the price y^{FP} for all T periods.

Under the joint inventory and pricing problem, the fixed-price policy sets initial inventory α^{FP} by solving

$$Q^{\mathsf{D}'}(T) \triangleq \max_{\alpha\geq 0} V^{\mathsf{D}',\alpha}(T) - c\alpha, \tag{S}$$

where $V^{\mathsf{D}',\alpha}$ is the deterministic revenue with initial inventory α and fixed-price policy $\mathbf{y}^{\mathsf{FP}}$.

Specifically,

$$V^{\mathsf{D}',\alpha}(T) \triangleq \sum_{t=1}^{T} x^{-1}\left(y^{\mathsf{FP}}\right)\lambda(n_{t-1}^{\mathsf{FP}},\alpha)y^{\mathsf{FP}}, \tag{24}$$

where $n_0^{\mathsf{FP}} = \alpha$ and $n_t^{\mathsf{FP}} = n_{t-1}^{\mathsf{FP}} - \lambda(n_{t-1}^{\mathsf{FP}},\alpha)y^{\mathsf{FP}}$ for all $t \leq T$. Then given α^{FP}, it sets $\mathbf{y}^{\mathsf{FP}}$ as the fixed-price policy just described with $\alpha = \alpha^{\mathsf{FP}}$. The fixed-price policy is outlined in Algorithm 4.

Algorithm 4: Setting the initial inventory and prices based on fixed-price policy.

```
procedure FIXED POLICY(T)
    α^FP ← optimal solution of (S)
    set N_0 = α^FP                                      ▷ set initial inventory
    set prices with FIXED PRICING(α^FP, T)

procedure FIXED PRICING(α, T)
    y^FP ← ȳ or ȳ^so based on cases in (23) for α
    for t ← 1 to T do
        set intensity y^FP by offering price x^{-1}(y^FP)   ▷ set current intensity (price)
```

We next state the main result, which describes the performance of the fixed-price policy compared to a T-price policy. The benefit of additional price-change opportunities is built on Proposition 6. Under the setting where the expected demand and the initial inventory are scaled by m, we denote the expected profit of the fixed-price policy $(m\alpha^{\mathsf{FP}}, \mathbf{y}^{\mathsf{FP}})$ as $Q^{\mathsf{FP}}(m,T)$. For any $\alpha \geq 0$, we denote $V^{\mathsf{FP},\alpha}(m,T)$ as the expected revenue under the stochastic model of the fixed-price policy $\mathbf{y}^{\mathsf{FP}}$ with initial inventory $m\alpha$, and $V^{*,\alpha}(m,T)$ as the expected revenue under the optimal T-price policy with initial inventory $m\alpha$. Proposition 6 indicates that the loss between the two pricing policies at least grows linearly. The proof of the following result is in Appendix D.2.

Proposition 6 (Profit loss of the fixed-price policy)**.** When $T \geq 2$, if the following conditions hold for a fixed $\alpha \geq 0$:

(i) $\frac{\partial}{\partial y} V^{\mathsf{D}}\left(T-1; \alpha - \lambda(\alpha,\alpha)y, \alpha\right)\Big|_{y=\bar{y}} \neq 0$, and

(ii) $\alpha \geq \sum_{t=1}^{T} \lambda(n_{t-1}^{\bar{y}}, \alpha)\bar{y}$,

then the expected revenue loss is

$$V^{*,\alpha}(m,T) - V^{\mathsf{FP},\alpha}(m,T) = \Omega(m(T-1)).$$

Moreover, if (i)–(ii) hold for $\alpha = \alpha^{\mathsf{FP}}$, then $Q^*(m,T) - Q^{\mathsf{FP}}(m,T) = \Omega(m(T-1))$.

Condition (i) of Proposition 6 implies the myopic optimal intensity $\bar{y}$ is not the optimal first period price for deterministic model $V^{\mathsf{D}}(T)$. Condition (ii) means that we have a sufficient amount of initial inventory if we use to set the price at $x^{-1}(\bar{y})$. Proposition 6 shows that both profit loss and revenue loss of a fixed price policy grows at least linearly in the market size m.

With Proposition 6 in hand, we can further explore the marginal benefit of additional pricing policy by comparing the T-price policy to the optimal T'-price policy with $T' < T$ in which the seller changes price T' times over T time periods. (assuming equal interval between price changes). We find that the revenue loss is at least in the order of $\Omega(m(T/T'-1))$. The statement is formally stated in Corollary 1. The reason is as follows. For a T'-price policy (T' prices for T periods) compared to a T-price policy (T prices for T periods), the number of periods between price changes in a T'-price policy is T/T'. Considering the final T/T' periods, the T'-price policy will set a constant price, whereas the T-price policy will set the T/T' prices that optimize the revenue-to-go of the last T/T' periods. This is the place to apply Proposition 6, which demonstrates the incremental benefit of adding an additional price change opportunity. We also explore this question numerically in the next section.

Corollary 1. Let $Q^*(m, T)$ denote the optimal expected profit when the seller can change the price T times for T periods. Let $Q^{T'}(m, T)$ denote the expected profit when the seller certainty equivalent optimally sets T' prices over the selling season of T periods. Then, the expected revenue loss is

$$Q^*(m, T) - Q^{T'}(m, T) = \Omega(m(T/T' - 1)).$$

6 Numerical Studies

In this section, we conduct several numerical experiments to demonstrate the performance of the certainty-equivalent policies (CE-OL and CE-CL).

6.1 Revenue loss of the certainty-equivalent policy

Following Example 3, we choose price sensitivity function $x(\pi) = e^{-\gamma\pi} - c_x$, where the small constant c_x is chosen for numerical stability purposes. We consider a case where the demand is influenced by both past purchases and inventory availability by setting the SIS function to be a mixture of the SIS functions in Examples 1 and 3, respectively. In particular,

$$\lambda(n, \alpha) = \left(w\lambda^{(1)}(n, \alpha) + (1 - w)\lambda^{(2)}(n, \alpha)\right) \Delta t, \tag{25}$$

where $\lambda^{(1)}(n, \alpha) = \left((n - \alpha^2 + 1)/N_r\right)^{\beta}$ (cf. (3a)) and $\lambda^{(2)}(n, \alpha) = (1 - (\alpha - n))\,(p + q(\alpha - n))$ (cf.(2)), and Δt is the constant length of each time period. (We include the constant Δt because later on, we examine the effect of changing Δt to change the number of price change opportunities within a fixed time.) Note that we modified (3a) so that $\lambda(n, \alpha)$ is jointly concave in (n, α). These modifications have no effect on the qualitative properties of the optimal prices in Smith and Agrawal (2017). Here $\lambda(n, \alpha)$ in (25) is jointly concave in (n, α). The parameters used in this example are $(p, q, N_r, \beta, \gamma, c_x, T, \Delta t) = (0.4, 0.6, 25, 0.6, 0.001, 0.01, 10, 2)$.

We next illustrate the performance of the CE policies on the demand pattern considered in (25). We set $w = 0.5$ in (25) so that both display and word-of-mouth effects are present. From the previous experiments, the CE policy sets initial inventory $\alpha^{\mathsf{CE}} = 0.84$. The dynamic pricing policy $\mathbf{y}^{\mathsf{CL}}$ is based on reoptimizing (**D**) in each period with updated inventory levels. The policy $\mathbf{y}^{\mathsf{OL}}$ does not reoptimize the revenue in each period but sets time-varying prices.

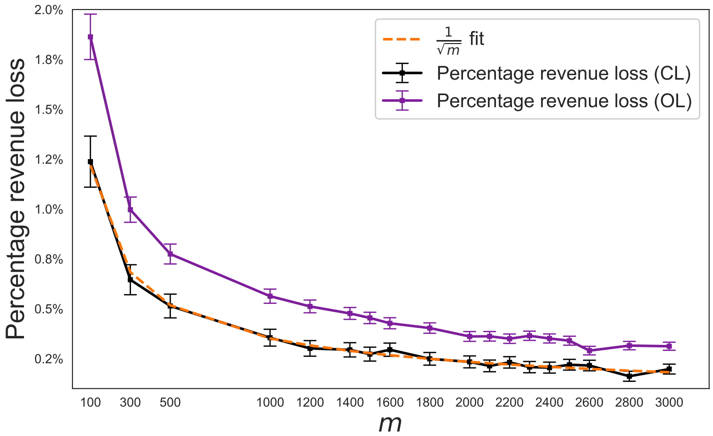

Figure 1: Upper bound on the percentage revenue loss of the certainty-equivalent policies against the optimal value of the stochastic problem. The fixed-price policy has a bound on percentage revenue loss that is at least 30% (not shown in graph).

We vary the inventory and demand scaling factor m from 100 to 3000, with discretizations shown in the horizontal axis of Figure 1. For each m, we randomly generate 2×10^4 demand sample paths following a bounded support Poisson distribution; we implement the dynamic pricing policies $\mathbf{y}^{\mathsf{OL}}$ and $\mathbf{y}^{\mathsf{CL}}$, and record the realized revenue on each path. The revenue averaged over the sample paths, which we denote by $\bar{V}^{\mathsf{OL}}(m,T)$ and $\bar{V}^{\mathsf{CL}}(m,T)$, are the approximations for the expected revenue of the certainty-equivalent policies, $V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{OL}}}(m,T)$ and $V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{CL}}}(m,T)$ respectively. We also note the 95% confidence intervals of this sample average.

Since the optimal revenue $V^*(m,T)$ is impossible to compute for problems with unknown distribution, we compute $V^{\mathsf{D}}(m,T)$ (which is an upper bound of $V^*(m,T)$) for comparison. Based on our sample approximation for $V^{\mathsf{OL}}(m,T)$ and $V^{\mathsf{CL}}(m,T)$ for each m, we compute an upper bound for the revenue losses of the CE-OL and CE-CL policies as $(V^{\mathsf{D}}(m,T) - \bar{V}^{\mathsf{OL}}(m,T))/V^{\mathsf{D}}(m,T)$ and $(V^{\mathsf{D}}(m,T) - \bar{V}^{\mathsf{CL}}(m,T))/V^{\mathsf{D}}(m,T)$, which are shown as the points in Figure 1. The figure also shows the 95% confidence intervals of the revenue loss bound. From Theorem 3, we know that the upper bound on the revenue loss is $\mathcal{O}(1/\sqrt{m})$, which is tightly traced by the $1/\sqrt{m}$ fit, shown with a dashed line in Figure 1. We further observe that the revenue losses by implementing both $\mathbf{y}^{\mathsf{OL}}$ and $\mathbf{y}^{\mathsf{CL}}$ are very small ($\sim 0.15\%$ when $m = 3000$). This implies that for a product with a scaling factor even as small as 100–3000 (small expected demand per period), the certainty-equivalent policies perform well. One may wonder how well the best fixed-price policy performs for the same problem. In all our examples, the fixed-price policy has a percentage revenue loss greater than or equal to 30% (we omit this from the figure to better highlight the difference between CE-OL, CE-CL, and the optimal policy).

6.2 The benefit of reoptimization in a non-asymptotic setting

In contrast to the open-loop policy CE-OL, the closed-loop policy CE-CL reoptimizes the deterministic model (**D**) in each price review period with updated state information. In our asymptotic analysis, we show that reoptimization does not reduce the convergence order of CE revenue

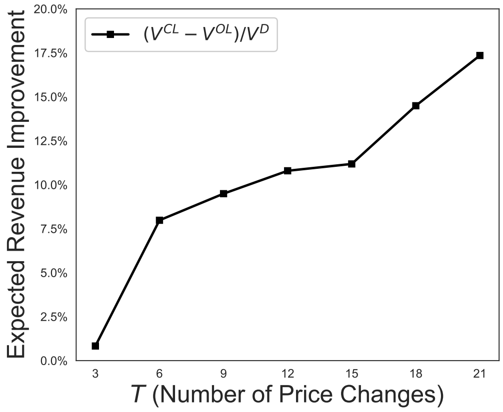

Figure 2: Value of resolving by increasing number of price changes (number of experiments to compute expectation: 20000)

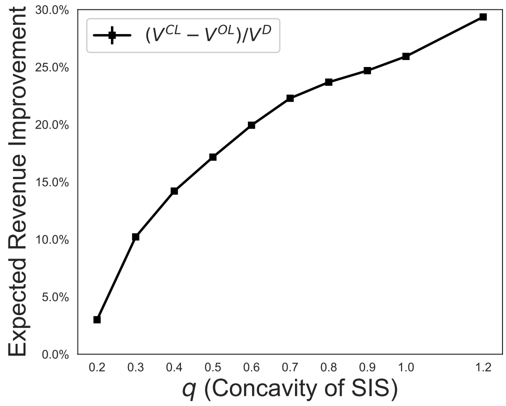

Figure 3: Value of resolving by increasing the concavity of the SIS function

loss. Through a numerical study comparing the two policies, we will examine the benefit of reoptimization in a non-asymptotic setting when $m = 1$.

Figure 2 shows how the gain from reoptimization is affected as the number of price changes T increases. In this example, demand follows a Poisson distribution and the Bass SIS function defined in (2) with p fixed at 0.01, when $q = 1.0$, and $k = 20$. The figure shows that more frequent reoptimization is beneficial as more opportunities to adjust prices reduces the probability of an early stock-out during the selling horizon and generates more revenue out of the remaining inventory. We note that the benefit of reoptimization has an increasing trend if there are more opportunities for changing prices.

Figure 3, on the other hand, shows how the gain from reoptimization changes by changing q while keeping everything else the same. Since $-\frac{\partial^2 \lambda}{\partial n^2} \propto q$, changing q is equivalent to changing the concavity of λ. Our example shows that the gain increases as the SIS function becomes more concave. This is because when the SIS function is highly non-linear and concave, the static CE-OL current price typically deviates more from the optimal policy. For instance, if the SIS function follows a Bass function, as defined in (2), the second-order derivative with respect to inventory decreases with q, where q is the imitation parameter in Bass terminology. This means that as q increases (i.e., more people imitate), the seller will lose significant revenue by not reoptimizing (**D**).

We next discuss the intuition on why a revenue gap between CE-OL and CE-CL exists. The closed-loop policy reoptimizes the price in each period so, given state information, its expected demand does not exceed the remaining inventory. Hence, the conditional expectation of its inventory path, $\mathbb{E}(N_t \mid \mathcal{F}_{t-1}) = N_{t-1} - \lambda(N_{t-1}, \alpha)\mathbf{y}^{\mathsf{CL}}(N_{t-1}, t)$, does not significantly deviate from its deterministic counterpart, $n_t^{\mathsf{D}} = n_{t-1}^{\mathsf{D}} - \lambda(n_{t-1}^{\mathsf{D}}, \alpha)y_t^{\mathsf{D}}$. In contrast, policy CE-OL does not guarantee that the conditional expectation of N_t is close to n_t^{D}. This is because, given the inventory state, the open-loop price can result in an expected demand that is greater than the inventory, so $\mathbb{E}(N_t \mid \mathcal{F}_{t-1}) \neq N_{t-1} - \lambda(N_{t-1}, \alpha)y_t^{\mathsf{OL}}$. This explains why the revenue loss relative to the deterministic upper bound is greater under CE-OL.

6.3 Revenue loss due to limited price changes

The certainty-equivalent policies we consider are discrete-time policies that assume the underlying demand is modeled as a discrete-time process. Hence, an interesting question to ask is: how much revenue can the limited price change policy lose if the true demand is a continuous-time process? To answer this question, we use one of the CE policies, CE-CL, to illustrate the performance. We run experiments on demand modeled as a continuous-time Markov chain with the state variable N^m, where $N^m = \alpha m, \alpha m - 1, ..., 0$. If n is the current inventory level, the transition rate is $\lambda(n, \alpha m)x(\pi)/\Delta t$, with $\lambda(n, \alpha m)$ given in (25). That is, conditional on current inventory level n, the probability of having one sale during a time period of length $o(t)$ is

$$\mathbb{P}\left(N^m_{t+o(t)} = n + 1 \,|N^m_t = n\right) = \lambda\left(n, \alpha m\right) x(\pi_t) o(t)$$

and there is $o(t)$ probability of having more than one sale during a time period of length $o(t)$.

To see the loss due to the discrete approximation, we experiment with different values for Δt, the length of time between price changes. We do this while keeping the total planning horizon length $\bar{T} = T\Delta t$ unchanged. In particular, the case when Δt approaches zero represents continuous price changes, which serves as a benchmark for the discrete-time model. For a given $(T, \Delta t)$ pair, we compute the CE-CL policy $(\alpha^{\mathsf{CE}}, \mathbf{y}^{\mathsf{CL}})$ and implement the discrete-time policy in 8×10^3 sample paths simulated from the continuous-time Markov chain process.

For the various values of T, Table 4 reports the average revenue (and 95% confidence intervals) of the certainty-equivalent policy normalized against the average revenue with $T = 45$ price changes (i.e., the continuous-time policy benchmark). Notice that we can see diminishing marginal returns when increasing the number of price changes. Consistent with Section 6.1, we observe a sharp increase in revenue when the number of price changes increases from 1 to 10. However, we observe that 10 price changes are almost as good as continuous price changes.

These results provide numerical evidence that a few price changes are good enough to capture the revenue from changing price continuously (which is very costly in practice) even with dependent demand. A small number of prices go a long way. Hence, our setting of limited price changes with state-dependent demand is a meaningful one to consider in practice. We believe the most important reason for this is the fact that the SIS function $\lambda(n, \alpha)$ is assumed to be jointly concave in (n, α) so that the demand rate is relatively "flat" compared to other convex forms. Moreover, because of the concavity of λ, in Lemma 4, we found that the deterministic optimal policy is Lipschitz continuous in the remaining inventory. This means the difference in the two policies is not too large when the inventory level changes, which implies the deterministic optimal policy is a relatively stable pricing policy. With the optimal price path to be relatively stable, a well-designed policy with one price change in the middle can have the ability to roughly trace the optimal path, which can recover most of the revenue. However, we note that such policy (piecewise constant pricing) is not asymptotically optimal in the face of a continuous-time dependent demand model.

Table 4: The expected revenue of the discrete-time policy normalized with the expected revenue of a continuous-time policy

T: Number of price changes	1	2	3	4	5	10	15	20	35	45
95 % CI upper bound	70.3%	95.7%	96.8%	97.9%	98.5%	99.4%	99.7%	99.9%	100.0%	100.0%
Expected normalized revenue	70.3%	95.6%	96.7%	97.9%	98.4%	99.4%	99.6%	99.8%	99.9%	100.0%
95 % CI lower bound	70.3%	95.6%	96.7%	97.9%	98.4%	99.3%	99.6%	99.8%	99.9%	100.0%

7 Conclusion

Certainty equivalent (CE) policies are widely used in practice because they are easy to compute and require a minimal amount of information. The performance guarantee of CE policies has been extensively studied in the literature under settings where demand is independent across periods and prices can be changed at any time. In contrast to the demand models studied in the previous literature on CE policies, our demand model is able to capture two distinct forces that critically influence future demand. The first force is that future demands are influenced by past sales. The word-of-mouth effect is an example of this force. The second force is that future demand is influenced by inventory availability. This force is often manifested in one of two forms: the scarcity effect (in the case of luxury or fad items) and the billboard effect, which is found in many markets today. Moreover, we consider a periodic review pricing policy, which is commonly practiced in reality. Periodic review requires consideration of stockouts, a complicating factor our the analysis.

We analyze two CE pricing policies: an open-loop CE policy (CE-OL) and a closed-loop CE policy (CE-CL). We show that as the scaling factor m increases, both CE pricing policies are asymptotically optimal with a revenue loss of $\mathcal{O}(\sqrt{m})$ when compared with the optimal policy. The revenue loss upper bound is tight as we show the revenue loss of a CE policy is lower bounded by $\Omega\left(\sqrt{m}\right)$. Our theoretical results show that when future demand is state-dependent, re-optimization may not necessarily improve the convergence order of CE revenue. In contrast, when demands are independent, one can expect that re-optimization reduces revenue loss to $\mathcal{O}(\log m)$. We then extend our results to the case where the seller chooses the initial inventory along with the price in each period. We also show that when demand depends on time, cumulative sales, and/or inventory availability, the asymptotic performance of CE policies does not change.

To further explore the difference between dynamic pricing under state-dependent demand and traditional demand assumptions used in the dynamic pricing literature, we also evaluate the performance of the static pricing policy (which was proven to be optimal in classical settings). We show that the revenue loss from static pricing can be huge, and it grows at least at the rate of a linear function when demand is dependent on cumulative sales and inventory.

An accompanying numerical study shows the performance and implementability of both CE pricing policies. We also show that the CE-CL policy performs close to optimality even in cases where the scaling factor is not large. Furthermore, we show that significant revenue improvement can be achieved with only a few price changes.

There are several future directions for our work. One is to extend the framework to the multi-product case, which involves specifying the dependency relationships among all products and across periods. All the results in our paper will hold as long as the dependency across products does not break the convexity of the pricing problem. Another extension is to consider strategic customers. The customers can strategically wait until there is a discount. Sapra et al. (2010) touches on this with the wait-list effect, where here it may be that a customer registers some interest in the product (follows on Twitter) but is waiting for a sale. A further direction is to incorporate learning into our model. In this paper, we assume that the conditional expectation of the demand is known. It may be possible to approximate the expectation using available data throughout the selling horizon. For nonparametric demand learning in our setting, one can employ a spline approximation because such approximation can approximate any continuous function on a compact set well. However, understanding the tradeoff between exploration and exploitation would be the key to designing the learning and pricing algorithm.

References

Aouad A, Ma W (2022) A nonparametric framework for online stochastic matching with correlated arrivals. *arXiv preprint arXiv:2208.02229* .

Bai Y, El Housni O, Jin B, Rusmevichientong P, Topaloglu H, Williamson DP (2023) Fluid approximations for revenue management under high-variance demand. *Management Science* .

Balachander S, Liu Y, Stock A (2009) An empirical analysis of scarcity strategies in the automobile industry. *Management Science* 55(10):1623–1637.

Balseiro SR, Besbes O, Pizarro D (2023) Survey of dynamic resource-constrained reward collection problems: Unified model and analysis. *Operations Research* .

Banker RD, Khosla I, Sinha KK (1998) Quality and competition. *Management Science* 44(9):1179–1192.

Bass FM (1969) A new product growth for model consumer durables. *Management Science* 15(5):215–227.

Bass FM, Krishnan TV, Jain DC (1994) Why the Bass model fits without decision variables. *Marketing Science* 13(3):203–223.

Besbes O, Zeevi A (2009) Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research* 57(6):1407–1420.

Bitran GR, Mondschein SV (1997) Periodic pricing of seasonal products in retailing. *Management Science* 43(1):64–79.

Boyd S, Vandenberghe L (2004) *Convex Optimization* (Cambridge University Press).

Bumpensanti P, Wang H (2020) A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science* 66(7):2993–3009.

Cachon GP, Gallino S, Olivares M (2018) Does adding inventory increase sales? Evidence of a scarcity effect in us automobile dealerships. *Management Science* 65(4):1469–1485.

Caro F, Gallien J (2012) Clearance pricing optimization for a fast-fashion retailer. *Operations Research* 60(6):1404–1422.

Chen F, Federgruen A, Zheng YS (2001) Near-optimal pricing and replenishment strategies for a retail/distribution system. *Operations Research* 49(6):839–853.

Chen Y, Farias VF (2013) Simple policies for dynamic pricing with imperfect forecasts. *Operations Research* 61(3):612–624.

Cheung WC, Simchi-Levi D, Wang H (2017) Dynamic pricing and demand learning with limited price experimentation. *Operations Research* 65(6):1722–1731.

Cui R, Zhang DJ, Bassamboo A (2019) Learning from inventory availability information: Field evidence from Amazon. *Management Science* 65(3):955–1453.

Datta T, Pal A (1990) A note on an inventory model with inventory-level-dependent demand rate. *Journal of the Operational Research Society* 41(10):971–975.

Deng S, Yano CA (2006) Joint production and pricing decisions with setup costs and capacity constraints. *Management Science* 52(5):741–756.

Feng Y, Gallego G (2000) Perishable asset revenue management with Markovian time dependent demand intensities. *Management Science* 46(7):941–956.

Gallego G, Van Ryzin G (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science* 40(8):999–1020.

Gallego G, Van Ryzin G (1997) A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research* 45(1):24–41.

Gerchak Y, Wang Y (1994) Periodic-review inventory models with inventory-level-dependent demand. *Naval Research Logistics* 41(1):99–116.

Geunes J, Romeijn HE, Taaffe K (2006) Requirements planning with pricing and order selection flexibility. *Operations Research* 54(2):394–401.

Jasin S (2014) Reoptimization and self-adjusting price control for network revenue management. *Operations Research* 62(5):1168–1178.

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

Jasin S, Kumar S (2013) Analysis of deterministic LP-based booking limit and bid price controls for revenue management. *Operations Research* 61(6):1312–1320.

Jiang J (2023) Constant approximation for network revenue management with markovian-correlated customer arrivals. *arXiv preprint arXiv:2305.05829* .

Jiang J, Ma W, Zhang J (2022) Degeneracy is ok: Logarithmic regret for network revenue management with indiscrete distributions. *arXiv preprint arXiv:2210.07996* .

Kopalle PK, Mela CF, Marsh L (1999) The dynamic effect of discounting on sales: Empirical analysis and normative pricing implications. *Marketing Science* 18(3):317–332.

Krishnamoorthy A, Prasad A, Sethi SP (2010) Optimal pricing and advertising in a durable-good duopoly. *European Journal of Operational Research* 200(2):486–497.

Krishnan TV, Bass FM, Jain DC (1999) Optimal pricing strategy for new products. *Management Science* 45(12):1650–1663.

Lariviere MA (2006) A note on probability distributions with increasing generalized failure rates. *Operations Research* 54(3):602–604.

Lei YM, Jasin S, Uichanco J, Vakhutinsky A (2021) Joint product framing (display, ranking, pricing) and order fulfillment under the mnl model for e-commerce retailers. *Manufacturing & Service Operations Management* Forthcoming.

Li W, Rusmevichientong P, Topaloglu H (2023) Revenue management with calendar-aware and dependent demands: Asymptotically tight fluid approximations. *Available at SSRN 4543277* .

Maglaras C, Meissner J (2006) Dynamic pricing strategies for multiproduct revenue management problems. *Manufacturing & Service Operations Management* 8(2):136–148.

Rajan A, Rakesh, Steinberg R (1992) Dynamic pricing and ordering decisions by a monopolist. *Management Science* 38(2):240–262.

Raman K, Chatterjee R (1995) Optimal monopolist pricing under demand uncertainty in dynamic markets. *Management Science* 41(1):144–162.

Reiman MI, Wang Q (2008) An asymptotically optimal policy for a quantity-based network revenue management problem. *Mathematics of Operations Research* 33(2):257–282.

Sapra A, Truong VA, Zhang RQ (2010) How much demand should be fulfilled? *Operations Research* 58(3):719–733.

Scarf H (1958) A min-max solution of an inventory problem. *Studies in the Mathematical Theory of Inventory and Production* .

Sethi SP, Prasad A, He X (2008) Optimal advertising and pricing in a new-product adoption model. *Journal of Optimization Theory and Applications* 139(2):351.

Shen W, Duenyas I, Kapuscinski R (2011) New product diffusion decisions under supply constraints. *Management Science* 57(10):1802–1810.

Shen W, Duenyas I, Kapuscinski R (2014) Optimal pricing, production, and inventory for new product diffusion under supply constraints. *Manufacturing & Service Operations Management* 16(1):28–45.

Smith SA, Achabal DD (1998) Clearance pricing and inventory policies for retail chains. *Management Science* 44(3):285–300.

Smith SA, Agrawal N (2017) Optimal markdown pricing and inventory allocation for retail chains with inventory dependent demand. *Manufacturing & Service Operations Management* 19(2):290–304.

Song I, Chintagunta PK (2003) A micromodel of new product adoption with heterogeneous and forward-looking consumers: application to the digital camera category. *Quantitative Marketing and Economics* 1:371–407.

Thomas J (1970) Price-production decisions with deterministic demand. *Management Science* 16(11):747–750.

Urban TL (2005) Inventory models with inventory-level-dependent demand: A comprehensive review and unifying theory. *European Journal of Operational Research* 162(3):792–804.

Urban TL, Baker R (1997) Optimal ordering and pricing policies in a single-period environment with multivariate demand and markdowns. *European Journal of Operational Research* 103(3):573–583.

Van Herpen E, Pieters R, Zeelenberg M (2009) When demand accelerates demand: Trailing the bandwagon. *Journal of Consumer Psychology* 19(3):302–312.

Wang Y, Gerchak Y (2001) Supply chain coordination when demand is shelf-space dependent. *Manufacturing & Service Operations Management* 3(1):82–87.

Wang Y, Wang H (2022) Constant regret resolving heuristics for price-based revenue management. *Operations Research* 70(6):3538–3557.

Wolfe HB (1968) A model for control of style merchandise. *Industrial Management Review* 9(2):69–82.

Yang N, Zhang R (2014) Dynamic pricing and inventory management under inventory-dependent demand. *Operations Research* 62(5):1077–1094.

Zhang M, Ahn HS, Uichanco J (2022) Data-driven pricing for a new product. *Operations Research* 70(2):847–866.

Ziya S, Ayhan H, Foley RD (2004) Relationships among three assumptions in revenue management. *Operations Research* 52(5):804–809.

Online Companion

Appendix A An interesting example

In the following example, we create a case of a CE policy not meeting the $\sqrt{m}$ criterion with Assumption 3 only. We will show that even though Assumption 3 is satisfied in the example, the CE policy may not be asymptotically optimal. We should note that in this example, the initial inventory is scaling down in the m-th problem but keeping $\alpha^m/\alpha = m$, so it is a slightly different asymptotic regime.

Example 5. Here, we give an example where CE policies fail when Assumption 3 is satisfied but Assumption 2 fails. We take $\lambda = 1, \lambda^m = m$, $x(\pi) = 1/\pi$, initial inventory is $\alpha = 1/m$, and the initial inventory in the m-th problem is $\alpha^m = 1$, demand is independently and identically distributed following a Poisson distribution, there is only one period $T = 1$, then this example still satisfies our Assumption 3 but Assumption 2 fails because $x(\pi) = 1/\pi$ is not even Lipschitz continuous and $\pi x(\pi)$ does not have a unique maximize.

In this example, any positive price no smaller than m is optimal in the CE problem. Let us take $\pi^{\mathsf{CE}} = m$ in the m-th problem. The expected revenue loss of this CE policy is not sublinear in m. The expected revenue loss is

$$
\begin{aligned}
V^{\mathsf{D}} - V^{\mathsf{CE}} &= m - \mathbb{E}\left[\pi^{\mathsf{CE}} \min\{D^m, 1\}\right] = m - m\mathbb{E}\left[D^m - (D^m - 1)^+\right] \\
&= m - m \cdot m \cdot \frac{1}{m} + m\mathbb{E}\left[(D^m - 1)^+\right] = m \cdot \sum_{d=1}^{\infty} (d-1)e^{-1}\frac{(1)^d}{(d)!} = \Theta(m/e).
\end{aligned}
$$

This example suggests that the mere enforcing of Assumption 3 is insufficient for guaranteeing $\mathcal{O}(\sqrt{m})$ regret. The other structural properties on λ are essential for deriving the bound.

Appendix B Proof of Lemma 1

Proof. We prove $r(y)$ is strictly concave in y first. Using the product and inverse differentiation rules, and the fact that $\pi = x^{-1}(y)$, yields

$$
\frac{d^2}{dy^2}[x^{-1}(y)y] = \frac{2 - \frac{x''(\pi)y}{x'(\pi)^2}}{x'(\pi)}.
$$

By Assumption 2(ii) the denominator is negative. Assumption 2(iii) implies, after taking derivatives, that $2 - \frac{x''(\pi)x(\pi)}{x'(\pi)^2} > 0$. Since $y_i \in [0, 1]$, this implies that the numerator is positive. Thus, $\frac{d^2}{dy^2}[x^{-1}(y)y] < 0$ and the strict concavity of $r(y)$ follows.

Besides the concavity of $r(y)$, the other properties are immediate from the relationships $y = x(\pi)$, $\rho(\pi) = r(y)$, the properties of x^{-1}, and Assumption 2(iv),(vii). □

Appendix C Section 3 proofs

C.1 Proof of Proposition 1

Proof. Any feasible solution to (**D**) is also feasible in ($\mathbf{D}^\dagger$), so $V^{\mathsf{D}}(T; u, \alpha) \le V^{\mathsf{D}^\dagger}(T; u, \alpha)$. To show "$\ge$", we will show that any feasible solution y to ($\mathbf{D}^\dagger$) where total demand exceeds inventory can be converted to a feasible solution with no stockout, and whose objective ($\mathbf{D}^\dagger$a) is at least as large as that of y.

Let $y = (y_1, y_2, \ldots, y_T)$ be any policy that has positive lost sales (n can be accordingly determined by y), i.e., $\lambda(n_{t-1}, \alpha)y_t > n_{t-1}$ for some period t. Let s be the index of the last period with lost sales.

We will modify policy y into a policy y' with one less period of lost sales, where the objective function ($\mathbf{D}^{\dagger}$a) under y' is no worse than that under y. More specifically, set $y'_s = \frac{n_{s-1}}{\lambda(n_{s-1},\alpha)}$, and $y'_t = y_t$ for all $t \neq s$. Note that y' is feasible to problem ($\mathbf{D}$) and $y'_s < y_s$.

The only difference between the objective value ($\mathbf{D}^{\dagger}$a) under y' and that under y is the revenue in period s. We have the difference to be

$$\underbrace{x^{-1}(y'_s)\min\left(\lambda(n_{s-1},\alpha)y'_s, n_{s-1}\right)}_{\text{revenue under } y'_s} - \underbrace{x^{-1}(y_s)\min\left(\lambda(n_{s-1},\alpha)y_s, n_{s-1}\right)}_{\text{revenue under } y_s}$$
$$= x^{-1}(y'_s)n_{s-1} - x^{-1}(y_s)n_{s-1} \geq 0$$

where the last inequality comes from the fact that $x^{-1}(\cdot)$ is a decreasing function by Assumption 2(ii). Hence, the objective of y' is no worse than that of y. We next modify the solution y' so that there is one less period with lost sales, and the objective is no worse. We do this until there are no more periods with lost sales. This completes our proof. $\square$

C.2 Proof of Proposition 2

Proof. (i) We first show that the objective function ($\mathbf{D}'$a) is jointly concave in d. To this end, we define the effective revenue function $r(y) := x^{-1}(y)y$, so the objective function ($\mathbf{D}'$a) is equivalent to

$$\sum_{t=1}^{T} \lambda(u - d_1 - d_2 - \ldots - d_{t-1}, \alpha) \cdot r\left(\frac{d_t}{\lambda(u - d_1 - d_2 - \ldots - d_{t-1}, \alpha)}\right). \tag{26}$$

To proceed, we require the following claim.

Claim 1. The function $(d', \lambda) \mapsto \lambda \cdot r\left(\frac{d'}{\lambda}\right)$ is strictly concave in (d', λ).

Claim 1 follows from Boyd and Vandenberghe (2004) page 39 (convexity of the perspective function).

We now show that each term in the summation of (26) is jointly concave in $(d_1, d_2, \ldots, d_T)$. Consider any $\theta \in [0,1]$, $d^1 = (d^1_1, d^1_2, \ldots, d^1_T)$ and $d^2 = (d^2_1, d^2_2, \ldots, d^2_T)$. We define the vector $\bar{d} \triangleq \theta d^1 + (1-\theta)d^2$, where $\bar{d}_t = \theta d^1_t + (1-\theta)d^2_t$.

Consider an arbitrary index t. Because $\lambda(n,\alpha)$ is jointly concave in (n,α) by Assumption 2(vi), then

$$\lambda(u - \bar{d}_1 - \bar{d}_2 - \ldots - \bar{d}_{t-1}, \alpha)$$
$$\geq \underbrace{\theta\lambda(u - d^1_1 - d^1_2 - \ldots - d^1_{t-1}, \alpha) + (1-\theta)\lambda(u - d^2_1 - d^2_2 - \ldots - d^2_{t-1}, \alpha)}_{\bar{\lambda}}. \tag{27}$$

From the definition of r, we have that

$$\lambda(u - \bar{d}_1 - \bar{d}_2 - \ldots - \bar{d}_{t-1}, \alpha) \cdot r\left(\frac{\bar{d}_t}{\lambda(u - \bar{d}_1 - \bar{d}_2 - \ldots - \bar{d}_{t-1}, \alpha)}\right)$$
$$= \bar{d}_t \cdot x^{-1}\left(\frac{\bar{d}_t}{\lambda(u - \bar{d}_1 - \ldots - \bar{d}_{t-1}, \alpha)}\right)$$
$$\geq \bar{d}_t \cdot x^{-1}\left(\frac{\bar{d}_t}{\bar{\lambda}}\right) = \bar{\lambda} \cdot r\left(\frac{\bar{d}_t}{\bar{\lambda}}\right)$$
$$\geq \theta\lambda(u - d^1_1 - \ldots - d^1_{t-1}, \alpha) \cdot r\left(\frac{d^1_t}{\lambda(u - d^1_1 - \ldots - d^1_{t-1}, \alpha)}\right)$$

$$+ (1-\theta)\lambda(u - d_1^2 - \ldots - d_{t-1}^2, \alpha) \cdot r\left(\frac{d_t^2}{\lambda(u - d_1^2 - \ldots - d_{t-1}^2, \alpha)}\right)$$

where the first inequality follows from the fact that x^{-1} is a monotone decreasing function and from (27). The second inequality follows Claim 1. Hence, this shows that each term in the summation (26) is jointly concave in $d = (d_1, \ldots, d_T)$. This proves that the objective function (**D**′a) is a jointly concave function in d.

We next show that the set of solutions d that satisfy constraints (**D**′b)–(**D**′c) is a convex set. To show this, we want to show that for any feasible $d^1 = (d_1^1, d_2^1, \ldots, d_T^1)$, $d^2 = (d_1^2, d_2^2, \ldots, d_T^2)$ and any $\theta \in [0,1]$, that $\bar{d} = \theta d^1 + (1-\theta)d^2$ is also feasible. Clearly, (**D**′b) is a linear constraint in d, so we only need to check that $\bar{d}_t \le \lambda(u - \bar{d}_1 - \ldots - \bar{d}_{t-1}, \alpha)$ for all t.

$$\begin{aligned}
\bar{d}_t &= \theta d_t^1 + (1-\theta)d_t^2 \\
&\le \theta\lambda(u - d_1^1 - \ldots - d_{t-1}^1, \alpha) + (1-\theta)\lambda(u - d_1^2 - \ldots - d_{t-1}^2, \alpha) \\
&\le \lambda(u - \bar{d}_1 - \ldots - \bar{d}_t, \alpha),
\end{aligned}$$

where the first inequality follows from the feasibility of d^1 and d^2, and the second inequality follows from (27). This completes the proof.

(ii) We prove the strict concavity of $V^{\mathsf{D}}(T; u, \alpha)$ through a reformulation of (**D**) using the transformation $d_t = \lambda(n_{t-1}, \alpha)y_t$ to yield:

$$\begin{aligned}
V^{\mathsf{D}}(T; u, \alpha) = \max_{n,d} \ & \sum_{t=1}^{T} x^{-1}\left(\frac{d_t}{\lambda(n_{t-1}, \alpha)}\right) \cdot d_t \\
\text{s.t.} \ & \sum_{t=1}^{T} d_t \le u \\
& n_t = n_{t-1} - d_t \quad \text{for all } t \ge 1 \\
& n_0 = u \\
& 0 \le d_t \le \lambda(n_{t-1}, \alpha) \quad \text{for all } t \ge 1.
\end{aligned} \tag{28}$$

For any $(u_1, \alpha_1) \ge 0$ and $(u_2, \alpha_2) \ge 0$, we denote the optimal solution of $V^{\mathsf{D}}(T; u_1, \alpha_1)$ and $V^{\mathsf{D}}(T; u_2, \alpha_2)$ by (n^1, d^1) and (n^2, d^2), respectively. We may assume, without loss of generality, that $(n^1, d^1) \ne (n^2, d^2)$. Given any $\theta \in (0,1)$, our goal is to construct a new solution from $(n^1, d^1), (n^2, d^2)$ that is feasible to (28) with $u = \bar{u} \triangleq \theta u_1 + (1-\theta)u_2$ and $\alpha = \bar{\alpha} \triangleq \theta\alpha_1 + (1-\theta)\alpha_2$, and whose objective value is strictly greater than $\theta V^{\mathsf{D}}(T; u_1, \alpha_1) + (1-\theta)V^{\mathsf{D}}(T; u_2, \alpha_2)$. Since $V^{\mathsf{D}}(T; \bar{u}, \bar{\alpha})$ is no smaller than the objective value of any feasible solution, $V^{\mathsf{D}}(T; \bar{u}, \bar{\alpha}) > \theta V^{\mathsf{D}}(T; u_1, \alpha_1) + (1-\theta)V^{\mathsf{D}}(T; u_2, \alpha_2)$. This proves the strict concavity of V^{D} in (u, α).

Set $\bar{n} \triangleq \theta n^1 + (1-\theta)n^2$ and $\bar{d} \triangleq \theta d^1 + (1-\theta)d^2$. It is easy to check that $(\bar{n}, \bar{d})$ is feasible to (28) with $u = \bar{u}$ and $\alpha = \bar{\alpha}$. This is because the set $\{(d_t, n_{t-1}), t \ge 1 : d_t - \lambda(n_{t-1}, \alpha) \le 0\}$ is convex due to the fact that $d_t - \lambda(n_{t-1}, \alpha)$ is convex in (d_t, n_{t-1}) and the intersection of convex sets are convex. It remains to show that this solution has a strictly better revenue than $\theta V^{\mathsf{D}}(T; u_1, \alpha_1) + (1-\theta)V^{\mathsf{D}}(T; u_2, \alpha_2)$. The revenue under $(\bar{n}, \bar{d})$ for period t is

$$g(\bar{d}_t, \bar{n}_t) \triangleq x^{-1}\left(\frac{\theta d_t^1 + (1-\theta)d_t^2}{\lambda(\theta n_t^1 + (1-\theta)n_t^2, \theta\alpha_1 + (1-\theta)\alpha_2)}\right) \cdot \left(\theta d_t^1 + (1-\theta)d_t^2\right).$$

Accordingly, our goal becomes showing

$$\sum_{t=1}^{T} g(\bar{d}_t, \bar{n}_t) > \theta \cdot \sum_{t=1}^{T} x^{-1}\left(\frac{d_t^1}{\lambda(n_t^1, \alpha_1)}\right) d_t^1 + (1-\theta) \cdot \sum_{t=1}^{T} x^{-1}\left(\frac{d_t^2}{\lambda(n_t^2, \alpha_2)}\right) d_t^2 \\ = \theta V^{\mathsf{D}}(\alpha_1, T) + (1-\theta) V^{\mathsf{D}}(\alpha_2, T). \tag{29}$$

In fact, we will show that there is a dominance of revenue in every period:

$$g(\bar{d}_t, \bar{n}_t) \geq \theta x^{-1}\left(\frac{d_t^1}{\lambda(n_t^1, \alpha_1)}\right) d_t^1 + (1-\theta) x^{-1}\left(\frac{d_t^2}{\lambda(n_t^2, \alpha_2)}\right) d_t^2 \tag{30}$$

and at least for one period the inequality is strict because otherwise $(n^1, d^1) = (n^2, d^2)$.
To show (30), we note that $g(d, n) = \lambda(n, \alpha) \cdot r\left(\frac{d}{\lambda(n,\alpha)}\right)$, where r is the effective revenue function $r(y) := x^{-1}(y)y$ defined in Appendix A.

We now show (30). We have

$$\begin{aligned}
g(\bar{d}_t, \bar{n}_t) &= \lambda(\bar{n}_t, \alpha) \cdot r\left(\frac{\bar{d}_t}{\lambda(\bar{n}_t, \alpha)}\right) \quad \text{(definition of } g) \\
&\geq x^{-1}\left(\frac{\theta d_t^1 + (1-\theta) d_t^2}{\theta \lambda(n_t^1, \alpha_1) + (1-\theta)\lambda(n_t^2, \alpha_2)}\right) \cdot \left(\theta d_t^1 + (1-\theta) d_t^2\right) \\
&= \left(\theta \lambda(n_t^1, \alpha_1) + (1-\theta)\lambda(n_t^2, \alpha_2)\right) \cdot r\left(\frac{\theta d_t^1 + (1-\theta) d_t^2}{\theta \lambda(n_t^1, \alpha_1) + (1-\theta)\lambda(n_t^2, \alpha_2)}\right) \quad \text{(definition of } r) \\
&\geq \theta \lambda(n_t^1, \alpha_1) r\left(\frac{d_t^1}{\lambda(n_t^1, \alpha_1)}\right) + (1-\theta)\lambda(n_t^2, \alpha_2) r\left(\frac{d_t^2}{\lambda(n_t^2, \alpha_2)}\right) \\
&= \theta x^{-1}\left(\frac{d_t^1}{\lambda(n_t^1, \alpha_1)}\right) d_t^1 + (1-\theta) x^{-1}\left(\frac{d_t^2}{\lambda(n_t^2, \alpha_2)}\right) d_t^2, \quad \text{(definition of } r)
\end{aligned}$$

where the first equality is from the definition of r, x^{-1} is a monotone decreasing function and because $\lambda(n, \alpha)$ is jointly concave in (n, α) by Assumption 2(vi), hence $\lambda(\bar{n}_t, \bar{\alpha}) \geq \theta \lambda(n_t^1, \alpha_1) + (1-\theta)\lambda(n_t^2, \alpha_2)$. The last inequality is from Claim 1. This establishes (30), which in turn yields (29). This completes the proof.

□

C.3 Proof of Proposition 3

Proof. We first show (**D**) has a unique solution. Then via the transformation in (**D′**), this implies that (**D′**) has a unique optimal solution.

We prove this result through a dynamic programming reformulation of the deterministic program (**D**). (Note that in practice this DP does not need to be solved to determine V^{D}, which can be found more efficiently using interior-point methods as we discuss in Section 3. This DP is only used for the purpose of analysis and proof.)

Fix α. For any $u \in [0, \alpha]$, consider the following dynamic programming counterpart of (**D**):

$$R^{\mathsf{D}}(u, T) = \max_{y} \quad x^{-1}(y)\lambda(u, \alpha)y + R^{\mathsf{D}}\left(u - \lambda(u, \alpha)y, T-1\right) \tag{31a}$$

$$\text{s.t.} \quad \lambda(u, \alpha)y \leq u, \tag{31b}$$

where the base case is $R^{\mathsf{D}}(u, 0) = 0$ for all $u \in [0, \alpha]$. Note that $V^{\mathsf{D}}(T; u, \alpha) = R^{\mathsf{D}}(u, T)$. Further, we can construct an optimal solution (**D**) by solving the dynamic programming equations (31). Hence, to show that (**D**) has a unique solution, we need to show that (31) has a unique solution. Since the feasible set of (31) is compact, to show that (31) has a unique solution, it suffices to show that the objective function,

$$R^{\mathsf{D},y}(u,T) \triangleq x^{-1}(y)\lambda(u,\alpha)y + R^{\mathsf{D}}\left(u-\lambda(u,\alpha)y, T-1\right) \tag{32}$$

is strictly concave in y.

Claim 2. $R^{\mathsf{D},y}(u,T)$ is strictly concave in y.

The first term of $R^{\mathsf{D},y}(u,T)$ is strictly concave in y from Lemma 1(i). To see that the second term is also concave, its second-order derivative with respect to y is

$$\lambda(u,\alpha)^2 \frac{\partial^2}{\partial u'^2} R^{\mathsf{D}}(u',T-1)\Big|_{u'=u-\lambda(u,\alpha)y} \leq 0,$$

where $|_{u'=u}$ means the term is evaluated at $u'=u$, and the inequality comes from Proposition 2(ii) and the fact that $R^{\mathsf{D}}(u',T-1) = V^{\mathsf{D}}(T;u',\alpha)$. □

C.4 Proof of Proposition 4

The proof requires the following lemma.

Lemma 5. Let (n,y) be a feasible solution to (**D**), where $y \neq 0$. If $y_i = 0$ for some index i, there exists a feasible solution (n',y') with $(n',y') \neq (n,y)$ and whose objective value is the same as (n,y).

Proof of Lemma 5. We define the following procedure to move $y_i = 0$ to the last period T to yield a solution (n',y') that gives the same objective value as (n,y).

Algorithm 5

procedure MOVE(i,n,y)
    $(n'_t = n_t, y'_t = y_t)$ for all $t \leq i-1$
    $(n'_t = n_{t+1}, y'_t = y_{t+1})$ for all $i \leq t \leq T-1$
    $(n'_T = n_T, y'_T = 0)$
    return (n',y')

Since $y \neq 0$, the new policy generated from MOVE(i,n,y) for an appropriately chosen i results in $(n',y') \neq (n,y)$. (This is not true if the only nonzero entry of y is the first index; in which case, we modify the move procedure so that $y_i = 0$ is moved to the first period.) It is easy to check that (n',y') is a feasible solution to (**D**) since (n,y) is feasible.

Finally, we show that (n',y') has the same objective value as (n,y). Notice that n' is constructed by shifting every n_t with $t \geq i+1$ to one index smaller. The ending period remaining inventory is $n'_T = n_T$. Hence,

$$\begin{aligned}\sum_{t=1}^{T} x^{-1}(y_t)\lambda(n_{t-1},\alpha)y_t &= \sum_{t=1}^{i-1} x^{-1}(y_t)\lambda(n_{t-1},\alpha)y_t + \sum_{t=i+1}^{T} x^{-1}(y_t)\lambda(n_{t-1},\alpha)y_t \\ &= \sum_{t=1}^{i-1} x^{-1}(y'_t)\lambda(n'_{t-1},\alpha)y'_t + \sum_{t=i}^{T-1} x^{-1}(y'_t)\lambda(n'_{t-1},\alpha)y'_t.\end{aligned}$$

Here, the first equality comes from $y_i = 0$. The second equality comes from how Algorithm 5 (MOVE(i)) constructs y'. □

Now we can proceed with the proof of the theorem.

Proof of Proposition 4. We denote the unique optimal solution to (**D**) by $(n^{\mathsf{D}}, y^{\mathsf{D}})$ where $n^{\mathsf{D}} = (n_0^{\mathsf{D}}, n_1^{\mathsf{D}}, \ldots, n_T^{\mathsf{D}})$ and $y^{\mathsf{D}} = (y_1^{\mathsf{D}}, \ldots, y_T^{\mathsf{D}})$. We show (**D**) has the following properties:

(i) the optimal solution is strictly positive (i.e., $d^{\mathsf{D}} > 0$), and

(ii) the remaining inventory n^{D} is a strictly decreasing sequence.

Then via the transformation in (**D′**), this implies that the optimal solution d^{D} to (**D′**) lies in the interior of the feasible set (i.e., $\lambda(u - d_1 - \ldots - d_{t-1}, \alpha) > d_t^{\mathsf{D}} > 0$).

We first claim that for any $u \in (0, \alpha]$, the optimal partial solution y^{D} of (**D**) is such that $y^{\mathsf{D}} \neq 0$. This is because the objective value of $y = 0$ is 0. However, the objective value for y' where $y_1' = u/\lambda(u, \alpha)$ and $y_i' = 0$ for $i \neq 1$ is $x^{-1}(u/\lambda(u, \alpha))\, u > 0$. Note that y' is feasible since y_1' is the intensity that depletes all remaining inventory u. Hence, $y = 0$ cannot be optimal, so $y^{\mathsf{D}} \neq 0$.

We prove that $y^{\mathsf{D}} > 0$ using contradiction. Assume there exists an i such that $y_i^{\mathsf{D}} = 0$. Then, according to Lemma 5, we can construct a different solution with the same objective value. This contradicts Proposition 3 that the optimal solution of (**D**) is unique. □

C.5 Strong duality of dynamic programming counterpart of (D)

For a fixed α, note that $R^{\mathsf{D}}(u, T)$ in (31) is the dynamic programming counterpart of (**D**). We next establish a strong duality result for the DP formulation. This result is used in later proofs, notably Proposition 5.

Lemma 6. Fix α. For any $u \in (0, \alpha]$,

$$R^{\mathsf{D}}(u, T) = \inf_{\mu \geq 0} L^{\mathsf{D},\mu}(u, T), \tag{33}$$

where, for any $\mu \geq 0$, $L^{\mathsf{D},\mu}(u, T)$ is defined as:

$$L^{\mathsf{D},\mu}(u, T) \triangleq \max_{y \in [0,1]} \left\{ x^{-1}(y)\lambda(u, \alpha)y + R^{\mathsf{D}}\left(u - \lambda(u, \alpha)y,\ T - 1\right) + \mu\left(u - \lambda(u, \alpha)y\right) \right\}. \tag{34}$$

Proof. We use Slater's condition for convex programming duality (see page 226 in Boyd and Vandenberghe 2004). Recall, to invoke the Slater condition, we need to show that (31) is a convex optimization problem with a feasible point that satisfies its constraints strictly. Observe that all the constraints in (31) are affine in y. The objective function is concave in y, as established in Claim 2. Hence, (31) is a convex optimization problem.

The next step is to demonstrate that there exists a feasible solution to (31) that satisfies the inequality constraint (31b) strictly. Notice that any $y \in (0, \min\{1, u/\lambda(u, \alpha)\})$ is strictly feasible to (31) because since $u > 0$ and with Assumption 2(v), $u/\lambda(u, \alpha) > 0$. Hence, Slater's condition implies (33) holds. □

C.6 Proof of Proposition 5

Proof. We first introduce the dynamic programming counterpart of (**D†**) for any $u \in [0, \alpha]$:

$$R^{\mathsf{D}_0}(u, T) \quad \triangleq \quad \max_{y \in [0,1]} \quad x^{-1}(y) \min\left(\lambda(u, \alpha)y, u\right) + R^{\mathsf{D}_0}\left([u - \lambda(u, \alpha)y]^+,\ T - 1\right).$$

Fix α. We will make use of mathematical induction on T to prove $R^*(u, T) \leq R^{\mathsf{D}_0}(u, T) = R^{\mathsf{D}}(u, T)$ for any $u \in [0, \alpha]$. If we are able to prove this, this proves the rest of the proposition since $V^*(T) = R^*(\alpha, T)$ and $V^{\mathsf{D}}(T) = R^{\mathsf{D}}(\alpha, T)$.

For the base case with $T = 1$, we define the optimal expected revenue $R^*(u, 1)$ for any given remaining inventory $u \leq \alpha$ as:

$$R^*(u,1) \triangleq \max_{y\in[0,1]} \mathbb{E}_{y,u}\left[x^{-1}(y)\min(D,u)\right] \tag{35a}$$

$$\leq \max_{y\in[0,1]} x^{-1}(y)\min(\mathbb{E}_{y,u}(D),u) \tag{35b}$$

$$= \max_{y\in[0,1]} x^{-1}(y)\min(\lambda(u,\alpha)y,u)\,. \tag{35c}$$

Here, (35b) comes from $\min(D, n)$ is a concave function of D and Jensen's inequality. (35c) comes from $\mathbb{E}_{y_0,u}(D) = \lambda(u, \alpha)y_0$. From the definition of $R^{\mathsf{D}_0}(u, 1)$, the right-hand side of (35c) is equal to $R^{\mathsf{D}_0}(u, 1)$. Therefore, we have that

$$R^*(u,1) \leq R^{\mathsf{D}_0}(u,1). \tag{36}$$

This finishes the base case of induction because $R^*(u, 1) \leq R^{\mathsf{D}_0}(u, 1) = R^{\mathsf{D}}(u, 1)$.

For the inductive step, assume that for any $T \leq T'$, we have $R^*(u, T) \leq R^{\mathsf{D}_0}(u, T)$ for any given $u \leq \alpha$. We prove $R^*(u, T' + 1) \leq R^{\mathsf{D}_0}(u, T' + 1)$ for all $u \leq \alpha$ to finish the inductive step.

Note that we can reformulate $R^*(u, T' + 1)$ as:

$$R^*(u,T'+1) = \max_{y\in[0,1]} \mathbb{E}_{y,u}\left[x^{-1}(y)\left(D-[D-u]^+\right) + R^*([u-D]^+,T')\right] \tag{37a}$$

$$\text{s.t. } \mathbb{E}_{y,u}(D) = \lambda(u,\alpha)y. \tag{37b}$$

Claim 3. The maximization problem (37) is feasible and $R^*(u, T' + 1)$ is bounded.

We know $y = 0$ is a feasible solution. Moreover, the objective function (37a) is bounded below by zero and bounded above by $x^{-1}(\bar{y})\lambda(u, \alpha)\bar{y} + \max_{u\in[0,1]} R^{\mathsf{D}_0}(u, T') < \infty$, where $\bar{y}$ is defined in Lemma 1(ii). This concludes the claim.

Now, consider any $y \in [0, 1]$ feasible to (37b). We denote its objective value (37a) as $V^y(u, T' + 1)$. Then for any γ, we have that

$$V^y(u,T'+1) \leq \mathbb{E}_{y,u}\left[x^{-1}(y)\left(D-[D-u]^+\right) + R^*([u-D]^+,T')\right] + \gamma\left(\mathbb{E}_{y,u}(D)-\lambda(u,\alpha)y\right) \tag{38a}$$

$$\leq \max_{y_0\in[0,1]} \mathbb{E}_{y_0,u}\Big[x^{-1}(y_0)\left(D-[D-u]^+\right) + R^*([u-D]^+,T') + \gamma\left(D-\lambda(u,\alpha)y_0\right)\Big] \tag{38b}$$

$$\leq \max_{y_0\in[0,1]} \mathbb{E}_{y_0,u}\Big[x^{-1}(y_0)\left(D-[D-u]^+\right) + R^{\mathsf{D}_0}([u-D]^+,T') + \gamma\left(D-\lambda(u,\alpha)y_0\right)\Big]. \tag{38c}$$

Here, (38c) comes from the inductive hypothesis. Since (38) is true for all feasible y, taking the supremum of $V^y(u, T' + 1)$ over $y \in [0, 1]$ satisfying (37b), we have that $R^*(u, T' + 1)$ is bounded above by (38c).

Note that (38c), and hence $R^*(u, T' + 1)$, is bounded above by

$$\max_{\substack{y_0\in[0,1],\\ d\in\Re}} \left\{x^{-1}(y_0)\left(d-[d-u]^+\right) + R^{\mathsf{D}_0}([u-d]^+,T') + \gamma\left(d-\lambda(u,\alpha)y_0\right)\right\}. \tag{39}$$

Note that (39) is an upper bound because d, being a decision variable that can take any value, results in a larger value than (38c). Since (39) is an upper bound to $V^y(u, T' + 1)$ for any values of γ, we take the infimum over all possible values resulting in the upper bound (40) as follows:

$$R^*(u,T'+1) \le \inf_{\gamma} \max_{\substack{y_0\in[0,1],\\ d\in\Re}} \left\{x^{-1}(y_0)\left(d-[d-u]^+\right) + R^{\mathsf{D}_0}([u-d]^+,T') + \gamma\left(d-\lambda(u,\alpha)y_0\right)\right\}. \qquad (40)$$

Next, we will prove that the right-hand side of (40) equals $R^{\mathsf{D}}(u,T'+1)$. Note that $\gamma=0$ is the solution to (40) because otherwise, d can be chosen such that the value of (40) is $+\infty$. Then, for the problem in (40), it suffices to restrict $d\le u$, since any $d>u$ does not improve the value of the objective function. Thus, we know for any $\mu\ge 0$, the right-hand side of (40) is upper bounded by

$$\inf_{\gamma} \max_{\substack{y_0\in[0,1],\\ d\le u}} \left\{x^{-1}(y_0)\left(d-[d-u]^+\right) + R^{\mathsf{D}_0}([u-d]^+,T') + \gamma\left(d-\lambda(u,\alpha)y_0\right) + \mu(u-d)\right\}. \qquad (41)$$

Because (41) is the upper bound of (40) for any $\mu\ge 0$, we can take the infimum of (41) and yield the final upper bound of (40) as follows

$$R^*(u,T'+1) \le \inf_{\gamma,\mu\ge 0} \max_{\substack{y_0\in[0,1],\\ d\le u}} \left\{x^{-1}(y_0)d + R^{\mathsf{D}_0}(u-d,T') + \gamma\left(d-\lambda(u,\alpha)y_0\right) + \mu(u-d)\right\}. \qquad (42)$$

Since R^{D} is equivalent to R^{D_0}, we observe that the right-hand side of (42) is the dual problem of R^{D} and according to Lemma 6, we can simplify (42) as

$$R^*(u,T'+1) \le \max_{y_0\in[0,1]} \left\{x^{-1}(y_0)\lambda(u,\alpha)y_0 + R^{\mathsf{D}}(u-\lambda(u,\alpha)y_0,T')\right\} = R^{\mathsf{D}}(u,T'+1).$$

This finishes our inductive step. □

C.7 Proof of Theorem 1

Proof. Let $\pi^u \triangleq \arg\max_{\pi\in[0,\bar\pi]} \pi x(\pi)$ and $\pi^{ic} \triangleq \arg\min_{\pi\in[0,\bar\pi]} |x(\pi)-\alpha/T|$ (Note that π^u,π^{ic} are the same as p^u and p^c in Besbes and Zeevi (2009)).

In this case, we have

$$\arg\max_{\pi\ge 0} \pi x(\pi) = \frac{a}{2b} \ge \frac{2\bar b\bar\pi}{2\bar b} = \bar\pi,$$

and

$$a \ge x(\pi) \ge a - b\bar\pi \ge \max\{2\bar b\bar\pi, \bar b\bar\pi + \alpha/T\} - \bar b\bar\pi \ge \alpha/T$$

Thus, $\pi^u=\bar\pi$ and $\pi^{ic}=\bar\pi$. Then, using the same argument as in the proof of Proposition 2 in Besbes and Zeevi (2009): "By Proposition 1 in Gallego and Van Ryzin (1994), any dynamic pricing policy will never price below π^u; hence, the optimal dynamic pricing policy of V^* is to price at π^u until the minimum of T and the time when inventory is depleted." This yields (in our notations, following the same arguments as in proof of Proposition 2 in Besbes and Zeevi (2009))

$$
\begin{aligned}
V^*(m,T) &= \pi^u \mathbb{E}\left[\min\left\{\sum_{t=1}^{T} D_t^m, m\alpha\right\}\right] \\
&= \pi^u \mathbb{E}\left[\sum_{t=1}^{T} D_t^m - \left(\sum_{t=1}^{T} D_t^m - m\alpha\right)^+\right] \\
&= \pi^u \sum_{t=1}^{T} m x(\pi^u) - \pi^u \mathbb{E}\left[\left(\sum_{t=1}^{T} D_t^m - m\alpha\right)^+\right] \\
&= V^{\mathsf{D}}(m,T) - \pi^u m\alpha e^{-m\alpha}\frac{(m\alpha)^{m\alpha}}{(m\alpha)!}.
\end{aligned}
$$

Therefore, sending m to ∞ and using Sterling's approximation, the following holds

$$
m^{1/2}(1 - V^*(m,T)/V^{\mathsf{D}}(m,T)) \to C_1.
$$

for some constant $C_1 > 0$.

Because we have $V^*(m,T) \geq V^{\mathbf{y}}(m,T)$ for any policy $\mathbf{y}$, then we have $V^{\mathsf{D}}(m,T) - V^{\mathbf{y}}(m,T) \geq \Omega(\sqrt{m})$. This concludes that in this constructed example (satisfying all our modeling assumptions), for any policy that uses the information of expectation, the ultimate lower bound of the revenue loss is in the order of $\sqrt{m}$. $\square$

C.8 Proof of Lemma 2

Proof. When demand is deterministic, Lemma 2 holds trivially.

When demand is not deterministic, we prove the lemma by induction. Defining $\bar{I}_t^m = \bar{N}_t^m/m$, let $\bar{I}^m = (\bar{I}_0^m, \ldots, \bar{I}_T^m)$ be the stochastic sequence of normalized inventory under policy $\mathbf{y}^{\mathsf{OL}}$. The base case is $t = 0$, where all policies start with $\bar{I}_0^m = \alpha = n_0^{\mathsf{D}}$, and hence $\lambda\left(\bar{I}_0^m, \alpha\right) = \lambda(n_0^{\mathsf{D}}, \alpha) = \lambda(\alpha, \alpha)$. Therefore, (17) and (18) hold for $t = 0$.

For the induction step, assume that (17) and (18) hold for $t-1$, i.e.,

$$
\mathbb{E}\left|\bar{I}_{t-1}^m - n_{t-1}^{\mathsf{D}}\right| \leq \Theta(1/\sqrt{m}) \tag{43}
$$

$$
\mathbb{E}\left|\lambda(\bar{I}_{t-1}^m, \alpha) - \lambda(n_{t-1}^{\mathsf{D}}, \alpha)\right| \leq \Theta(1/\sqrt{m}). \tag{44}
$$

We prove that both properties hold for t.

To prove (17) for t, notice that by adding and subtracting $\mathbb{E}\left(\bar{I}_t^m\right)$,

$$
\mathbb{E}\left|\bar{I}_t^m - n_t^{\mathsf{D}}\right| = \mathbb{E}\left|\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right) + \mathbb{E}\left(\bar{I}_t^m\right) - n_t^{\mathsf{D}}\right| \leq \mathbb{E}\left|\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right| + \left|\mathbb{E}\left(\bar{I}_t^m\right) - n_t^{\mathsf{D}}\right|. \tag{45}
$$

We will show that both terms in (45) are $\mathcal{O}(1/\sqrt{m})$.

Consider the first term, $\mathbb{E}\left|\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right|$.

Claim 4. The following holds:

$$
\mathbb{E}\left|\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right| \leq \Theta(1/\sqrt{m}).
$$

Proof of Claim 4. Note that

$$
\mathbb{E}\left|\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right| = \mathbb{E}\left[\sqrt{\left(\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right)^2}\right].
$$

Since $\sqrt{x}$ is a concave function of x, applying Jensen's inequality, we have

$$\mathbb{E}\left[\sqrt{\left(\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right)^2}\right] \le \sqrt{\mathbb{E}\left(\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right)^2} = \sqrt{\mathrm{Var}\left(\bar{I}_t^m\right)}.$$

Next, we will prove $\mathrm{Var}\left(\bar{I}_t^m\right) \le \Theta(1/m)$ and $\mathrm{Var}\left(\lambda(\bar{I}_t^m, \alpha)\right) \le \Theta(1/m)$ by induction on t. When $t = 0$, the result holds as we have $\mathrm{Var}\left(\bar{I}_0^m\right) = \mathrm{Var}(\alpha) = 0 \le \Theta(1/m)$ and $\mathrm{Var}\left(\lambda(\bar{I}_0^m, \alpha)\right) = \mathrm{Var}(\lambda(\alpha, \alpha)) = 0 \le \Theta(1/m)$.

When $t \ge 1$, we assume $\mathrm{Var}\left(\bar{I}_{t-1}^m\right) \le \Theta(1/m)$ and $\mathrm{Var}\left(\lambda(\bar{I}_{t-1}^m, \alpha)\right) \le \Theta(1/m)$ holds.

By definition, we have $\bar{I}_t^m = \left[\bar{I}_{t-1}^m - \frac{D_t^m}{m}\right]^+ = \frac{1}{m}\left[\bar{N}_{t-1}^m - D_t^m\right]^+$. Moreover, according to the Cauchy-Schwarz inequality, we have $\left|\mathrm{Cov}\left(N_{t-1}^m, D_t^m\right)\right| \le \sqrt{\mathrm{Var}\left(D_t^m\right)\mathrm{Var}\left(N_{t-1}^m\right)}$. Then, we know $\left|\mathrm{Cov}\left(N_{t-1}^m, D_t^m\right)\right| \le \max\left\{\mathrm{Var}\left(D_t^m\right), \mathrm{Var}\left(N_{t-1}^m\right)\right\}$. Therefore, we have

$$\begin{aligned}
\mathrm{Var}\left(\bar{I}_t^m\right) &= \mathrm{Var}\left(\frac{1}{m}\left[\bar{N}_{t-1}^m - D_t^m\right]^+\right) = \frac{1}{m^2}\mathrm{Var}\left(\left[\bar{N}_{t-1}^m - D_t^m\right]^+\right) \\
&\le \frac{1}{m^2}\mathrm{Var}\left(D_t^m\right) + \frac{1}{m^2}\mathrm{Var}\left(N_{t-1}^m\right) + \frac{1}{m^2}\left|\mathrm{Cov}\left(N_{t-1}^m, D_t^m\right)\right| \\
&\le \Theta\left(\frac{1}{m^2}\mathrm{Var}\left(D_t^m\right) + \frac{1}{m^2}\mathrm{Var}\left(N_{t-1}^m\right)\right) \\
&\le \Theta\left(\frac{1}{m^2}\mathrm{Var}\left(D_t^m\right)\right) + \Theta(1/m).
\end{aligned}$$

The last inequality here is due to the inductive assumption.

Conditioning on $\mathcal{F}_{t-1}$, we have

$$\begin{aligned}
\mathrm{Var}\left(D_t^m\right) &= \mathbb{E}\left[\mathrm{Var}\left(D_t^m|\mathcal{F}_{t-1}\right)\right] + \mathrm{Var}\left[\mathbb{E}\left(D_t^m|\mathcal{F}_{t-1}\right)\right] \\
&\le \mathbb{E}\left(\sigma\mathbb{E}\left(D_t^m|\mathcal{F}_{t-1}\right)\right) + \mathrm{Var}\left[m\lambda(\bar{I}_{t-1}^m, \alpha)y_t^{\mathsf{OL}}\right] \\
&\le \sigma m\mathbb{E}\left(\lambda(\bar{I}_{t-1}^m, \alpha)y_t^{\mathsf{OL}}\right) + \Theta(m) \\
&\le \sigma m\left(\lambda(n_{t-1}^{\mathsf{D}}, \alpha)y_t^{\mathsf{OL}} + \Theta(1/\sqrt{m})\right) + \Theta(m) = \Theta(m).
\end{aligned} \tag{46}$$

The first inequality is due to Assumption 3 and the definition of $\mathbb{E}\left(D_t^m|\mathcal{F}_{t-1}\right)$. The second inequalities come from the inductive assumption that $\mathrm{Var}\left(\lambda(\bar{I}_{t-1}^m, \alpha)\right) \le \Theta(1/m)$ holds. The third inequalities come from the inductive assumption (44).

Therefore, we have

$$\mathrm{Var}\left(\bar{I}_t^m\right) \le \frac{1}{m^2}\mathrm{Var}\left(D_t^m\right) + \Theta(1/m) = \Theta(1/m).$$

For the variance of $\lambda(\bar{I}_t^m, \alpha)$, we note that because $\lambda(\cdot, \cdot)$ is a continuously differentiable function, then it is Lipschitz continuous. Thus, there exists a constant $C_\lambda = \max_{0\le n\le\alpha}\left|\frac{\partial\lambda(n,\alpha)}{\partial n}\right|$,

$$\begin{aligned}
&\mathrm{Var}\left(\lambda(\bar{I}_t^m, \alpha)\right) = \mathrm{Var}\left(\lambda(\bar{I}_t^m, \alpha) - \lambda\left(\mathbb{E}\left(\bar{I}_t^m\right), \alpha\right)\right) \le \mathbb{E}\left(\left[\lambda(\bar{I}_t^m, \alpha) - \lambda\left(\mathbb{E}\left(\bar{I}_t^m\right), \alpha\right)\right]^2\right) \\
&\le C_\lambda^2\mathbb{E}\left(\left[\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right]^2\right) = C_\lambda^2\mathrm{Var}(\bar{I}_t^m) \le \Theta(1/m).
\end{aligned}$$

This finishes the proof. □

Because $\bar{I}_t^m = \bar{N}_t^m/m$, the first term on the RHS of (45) is $\mathcal{O}(m^{-\frac{1}{2}})$.

For the second term in (45), we want to bound the difference between $\mathbb{E}\left(\bar{I}_t^m\right)$ and n_t^{D}. From the definition of $\bar{I}_t^m$, we know

$$\mathbb{E}\left(\bar{I}_t^m \mid \mathcal{F}_{t-1}\right) = \mathbb{E}\left(\left[\bar{I}_{t-1}^m - \frac{D_t^m}{m}\right]^+ \mid \mathcal{F}_{t-1}\right) = \frac{1}{m}\mathbb{E}\left(\left[\bar{N}_{t-1}^m - D_t^m\right]^+ \mid \mathcal{F}_{t-1}\right).$$

A well-known result by Scarf (1958) is that for any random variable X with mean μ and standard deviation σ, the following holds

$$\mathbb{E}([a-X]^+) \leq \frac{1}{2}\left(\sqrt{\sigma^2+(\mu-a)^2}-(\mu-a)\right), \tag{47}$$

and another version is

$$\mathbb{E}([X-a]^+) \leq \frac{1}{2}\left(\sqrt{\sigma^2+(\mu-a)^2}-(a-\mu)\right). \tag{48}$$

Note $\mathbb{E}(D_t^m \mid \mathcal{F}_{t-1}) = \lambda^m\left(\bar{N}_{t-1}^m, \alpha m\right) y_t^{\mathsf{OL}}$ and, by Assumption 3, $\mathrm{Var}(D_t^m \mid \mathcal{F}_{t-1}) \leq \sigma\lambda^m\left(\bar{N}_{t-1}^m, \alpha m\right) y_t^{\mathsf{OL}}$. Since $\bar{N}_{t-1}^m$ is not random when conditioning on the filtration $\mathcal{F}_{t-1}$, and from (15) we have

$$\begin{aligned}
\mathbb{E}\left(\bar{I}_t^m \mid \mathcal{F}_{t-1}\right) &\leq \frac{1}{2}\left(\sqrt{\frac{\sigma\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}}{m}+(\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}-\bar{I}_{t-1}^m)^2}-\left(\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}-\bar{I}_{t-1}^m\right)\right)\\
&\leq \frac{1}{2}\left(\sqrt{\frac{\sigma\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}}{m}}+\left|\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}-\bar{I}_{t-1}^m\right|-\left(\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}-\bar{I}_{t-1}^m\right)\right),
\end{aligned}$$

where the inequality is because $\bar{I}_t^m = (I_{t-1}^m - D_t^m/m)^+$.

Taking the expectation on both sides conditional on $\mathcal{F}_0$, we get

$$\begin{aligned}
&\mathbb{E}\left(\bar{I}_t^m \mid \mathcal{F}_0\right)\\
&\leq \frac{1}{2}\mathbb{E}\left[\sqrt{\frac{\sigma\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}}{m}}+\left|\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}-\bar{I}_{t-1}^m\right|-\left(\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}-\bar{I}_{t-1}^m\right) \mid \mathcal{F}_0\right]\\
&\leq \frac{1}{2}\left(\Theta(m^{-\frac{1}{2}})+n_{t-1}^{\mathsf{D}}-\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}+n_{t-1}^{\mathsf{D}}-\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}\right)\\
&= \Theta(m^{-\frac{1}{2}})+n_{t-1}^{\mathsf{D}}-\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}.
\end{aligned} \tag{49}$$

The last inequality comes from the inductive hypotheses (43),(44). In addition to this upper bound, we know that $\mathbb{E}(\bar{I}_t^m \mid \mathcal{F}_0) = \mathbb{E}\left(\left[\bar{I}_{t-1}^m - D_t^m/m\right]^+\right)$ is lower bounded by

$$\mathbb{E}\left(\mathbb{E}\left(\bar{I}_{t-1}^m - \frac{D_t^m}{m} \mid \mathcal{F}_{t-1}\right)\right) = \mathbb{E}\left(\bar{I}_{t-1}^m - \lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}\right) \geq n_{t-1}^{\mathsf{D}} - \lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}} - \Theta(m^{-\frac{1}{2}}), \tag{50}$$

where the equality follows from (15). The inequality is from the inductive hypothesis. Hence, (49) and (50) imply that

$$\left|\mathbb{E}(\bar{I}_t^m) - n_t^{\mathsf{D}}\right| = \left|\mathbb{E}(\bar{I}_t^m) - n_{t-1}^{\mathsf{D}} + \lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}\right| = \mathcal{O}(m^{-\frac{1}{2}})$$

Therefore, we can conclude that the RHS two terms of (45) are both bounded by $\mathcal{O}(m^{-\frac{1}{2}})$, thus giving us (17) for all t. For a given t, (18) follows by the Lipschitz continuity of λ and (17):

$$\mathbb{E}\left|\lambda\left(\bar{I}_t^m,\alpha\right) - \lambda(n_t^{\mathsf{D}},\alpha)\right| \leq C_\lambda\mathbb{E}\left|\bar{I}_t^m - n_t^{\mathsf{D}}\right| = \mathcal{O}\left(m^{-\frac{1}{2}}\right).$$

This concludes the proof. □

C.9 Proof of Lemma 3

Proof. Let $(n^{\mathsf{D}}, y^{\mathsf{D}})$ be the optimal solution of (**D**) with initial inventory α and $u = \alpha$. We can easily check that, because $\lambda^m(mn, m\alpha) = m\lambda(n, \alpha)$ for any $n \in [0, \alpha]$ because of (15), (**D**) with an initial inventory $m\alpha$ will have an optimal solution $(mn^{\mathsf{D}}, y^{\mathsf{D}})$. Therefore, $V^{\mathsf{D}}(m, T) = \sum_{t=1}^{T} x^{-1}(y_t^{\mathsf{D}}) m\lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}}$. By factoring out m, we can write the LHS of (19) as

$$
\begin{aligned}
& m \left| \mathbb{E}\left[\sum_{t=1}^{T} \left(x^{-1}\left(y_t^{\mathsf{OL}}\right) \lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha \right) y_t^{\mathsf{OL}} - x^{-1}(y_t^{\mathsf{D}}) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}} \right) \right] \right| \\
& \leq m\mathbb{E} \left| \sum_{t=1}^{T} \left(x^{-1}\left(y_t^{\mathsf{OL}}\right) \lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha \right) y_t^{\mathsf{OL}} - x^{-1}(y_t^{\mathsf{D}}) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}} \right) \right| \\
& \leq m \sum_{t=1}^{T} \mathbb{E} \left| x^{-1}\left(y_t^{\mathsf{OL}}\right) \lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha \right) y_t^{\mathsf{OL}} - x^{-1}(y_t^{\mathsf{D}}) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}} \right| \\
& = m \sum_{t=1}^{T} x^{-1}(y_t^{\mathsf{D}}) y_t^{\mathsf{D}} \mathbb{E} \left| \lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha \right) - \lambda(n_{t-1}^{\mathsf{D}}, \alpha) \right|. \qquad (51)
\end{aligned}
$$

Here, the first inequality comes from $|\mathbb{E}X| \leq \mathbb{E}|X|$ as a result of Jensen's inequality. The second inequality comes from the triangle inequality and the linearity of expectation. To prove the proposition, since T is a finite number, it is sufficient to show each term inside the summation of (51) is $\mathcal{O}(m^{-\frac{1}{2}})$.

This is true because, from (18) of Lemma 2, we know for any t,

$$\mathbb{E} \left| \lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha \right) - \lambda(n_{t-1}^{\mathsf{D}}, \alpha) \right| = \mathcal{O}(m^{-\frac{1}{2}}).$$

This concludes the proof. □

C.10 Proof of Theorem 2

Proof. First, we note that $V^*(m, T)$ is greater or equal to the revenue from a single-price policy and so is strictly positive. To prove the theorem, it is sufficient to show that

$$1 - \frac{V^{\mathsf{OL}}(m, T)}{V^{\mathsf{D}}(m, T)} \leq 1 - (1 - k) \left(1 - \frac{C}{2} \sqrt{\frac{\sigma}{m}} - k \right), \qquad (52)$$

where $k = \Theta(1/\sqrt{m})$ and C is some constant that is independent of m.

Let $\bar{N} = (\bar{N}_0^m, \ldots, \bar{N}_T^m)$ be the stochastic sequence of remaining inventories under $\mathbf{y}^{\mathsf{OL}}$ and define $\bar{I}_t^m \triangleq \bar{N}_t^m / m$. From (6), we have

$$V^{\mathsf{OL}}(m, T) = \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{E}\left[x^{-1}\left(y_t^{\mathsf{OL}}\right) \left(D_t^m - [D_t^m - \bar{N}_{t-1}^m]^+ \right) \mid \mathcal{F}_{t-1} \right] \right]. \qquad (53)$$

Note that $\bar{N}_{t-1}^m$ and $\bar{I}_{t-1}^m$ are not random when conditioning on the filtration $\mathcal{F}_{t-1}$. Furthermore, we have $\mathbb{E}(D_t^m \mid \mathcal{F}_{t-1}) = m\lambda(\bar{I}_{t-1}^m, \alpha) y_t^{\mathsf{OL}}$ and, by Assumption 3, $\mathrm{Var}(D_t^m \mid \mathcal{F}_{t-1}) \leq \sigma m\lambda(\bar{I}_{t-1}^m, \alpha) y_t^{\mathsf{OL}}$. Hence, by applying the Scarf bound (48) and from (15), which yields $\mathbb{E}(D_t^m \mid \mathcal{F}_{t-1}) = m\lambda(\bar{I}_{t-1}^m, \alpha) y_t^{\mathsf{OL}}$, we get

$$
\begin{aligned}
&\mathbb{E}\left[\left[D_t^m - \bar{N}_{t-1}^m\right]^+ \mid \mathcal{F}_{t-1}\right] \\
&\leq \frac{\sqrt{\operatorname{Var}(D_t^m \mid \mathcal{F}_{t-1}) + \left(\bar{N}_{t-1}^m - \mathbb{E}(D_t^m \mid \mathcal{F}_{t-1})\right)^2}}{2} - \frac{\left(\bar{N}_{t-1}^m - \mathbb{E}(D_t^m \mid \mathcal{F}_{t-1})\right)}{2} \\
&\leq \frac{\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}} + \left(\bar{N}_{t-1}^m - m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}\right)^2}}{2} - \frac{\left(\bar{N}_{t-1}^m - m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}\right)}{2} \\
&\leq \frac{1}{2}\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}} + \frac{1}{2}\left|\bar{N}_{t-1}^m - m\lambda(\bar{I}_{t-1}^m, \alpha) y_t^{\mathsf{OL}}\right| - \frac{1}{2}\left(\bar{N}_{t-1}^m - m\lambda(\bar{I}_{t-1}^m, \alpha) y_t^{\mathsf{OL}}\right). \qquad (54)
\end{aligned}
$$

Taking the expectation conditioning on $\mathcal{F}_0$ on both sides of (54), we have

$$
\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[\left[D_t^m - \bar{N}_{t-1}^m\right]^+ \mid \mathcal{F}_{t-1}\right] \mid \mathcal{F}_0\right] \\
&\leq \mathbb{E}\left[\frac{1}{2}\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}}\right] + \mathcal{O}(\sqrt{m}) + \frac{1}{2}\left|m n_{t-1}^{\mathsf{D}} - m\lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}}\right| - \frac{1}{2}\left[m n_{t-1}^{\mathsf{D}} - m\lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}}\right] \\
&= \mathbb{E}\left[\frac{1}{2}\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}}\right] + \mathcal{O}\left(\sqrt{m}\right). \qquad (55)
\end{aligned}
$$

Here, the first inequality comes from Lemma 2 and since $y_t^{\mathsf{OL}} = y_t^{\mathsf{D}}$ for all t. The equality is because $n_t^{\mathsf{D}} = n_{t-1}^{\mathsf{D}} - \lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}}$ due to constraint (**D**c), and $n_t^{\mathsf{D}} \geq 0$ due to the no-stockout constraint (**D**c).

Therefore, using (15) and plugging (55) into the RHS of (53) yields

$$
\begin{aligned}
V^{\mathsf{OL}}(m,T) &\geq \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\left(m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}} - \frac{1}{2}\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}}\right)\right] - \mathcal{O}(\sqrt{m}) \\
&= \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right) m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}\right] - \frac{1}{2}\sqrt{\sigma m}\,\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\sqrt{\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}}\right] - \mathcal{O}(\sqrt{m}) \\
&= \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right) m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}\right] \\
&\quad \times \left(1 - \frac{1}{2}\sqrt{\frac{\sigma}{m}}\underbrace{\frac{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\sqrt{\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}}\right]}{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}\right]}}_{(**)} - \mathcal{O}\left(1/\sqrt{m}\right)\right). \qquad (56)
\end{aligned}
$$

We get the first equality by multiplying x^{-1} term inside. The second equality comes from pulling out the first expectation term.

We first derive a lower bound for the first term in (56). Note that $\mathbf{y}^{\mathsf{OL}}$ does not scale with m since it is constructed from solutions of (**D**), which do not depend on m. From Lemma 3, we know that the difference between the first term in (56) and $V^{\mathsf{D}}(m,T)$ scales in $\mathcal{O}(\sqrt{m})$. This is slower than the speed of scaling $\Theta(m)$ of $V^{\mathsf{D}}(m,T)$. Hence,

$$
\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right) m\lambda\left(\bar{I}_{t-1}^m, \alpha\right) y_t^{\mathsf{OL}}\right) \geq V^{\mathsf{D}}(m,T)(1-k), \qquad (57)
$$

where $k = \Theta(m^{-\frac{1}{2}})$.

Next, we derive an upper bound for the term $(**)$, which results in a lower bound for the middle term in (56). Note that from the Cauchy-Swartz inequality, the numerator of $(**)$ is bounded above by

$$\mathbb{E}\left[\sqrt{\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}}\sqrt{\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)}\right] \leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}}\right]\sqrt{Tx^{-1}(0)}$$
$$\leq \sqrt{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}\right]}\sqrt{Tx^{-1}(0)},$$

where the first inequality comes from Assumption 2(ii), and the last inequality comes from Jensen's inequality and the fact that $\sqrt{z}$ is a concave function. Hence,

$$(**) \leq \sqrt{\frac{Tx^{-1}(0)}{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}\right]}} \leq \sqrt{\frac{Tx^{-1}(0)}{V^{\mathsf{D}}(\alpha,T)(1-k)}},$$

where the last inequality comes from (57).

Since $\Theta(m^{-\frac{1}{2}})$ decreases as m grows, we know there exists some constant $\Theta(1)$, unaffected by m, such that $\Theta(m^{-\frac{1}{2}}) \leq \Theta(1)$. Therefore, we know

$$\sqrt{\frac{1}{1-k}} = \sqrt{\frac{1}{1-\Theta(m^{-\frac{1}{2}})}} \leq \sqrt{\frac{1}{1-\Theta(1)}} = \Theta(1).$$

Hence, we have that

$$(**) \leq \sqrt{\frac{Tx^{-1}(0)}{V^{\mathsf{D}}(T)}}\Theta(1) \triangleq C. \tag{58}$$

Finally, we take (57) and (58) into (56), resulting in

$$V^{\mathsf{OL}}(m,T) \geq V^{\mathsf{D}}(m,T)(1-\mathcal{O}(1/\sqrt{m}))\left(1-\frac{1}{2}\sqrt{\frac{\sigma}{m}}C-\mathcal{O}(1/\sqrt{m})\right).$$

This completes the proof. □

C.11 Proof of Lemma 4

Proof. We prove the lemma by showing that $\mathbf{y}^{\mathsf{CL}}(n,t)$ has a bounded derivative with respect to n for $n \in [0,\alpha]$ because

$$\left|\mathbf{y}^{\mathsf{CL}}(n,t)-\mathbf{y}^{\mathsf{CL}}(n',t)\right| = \left|\int_{n'}^{n}\frac{\partial \mathbf{y}^{\mathsf{CL}}(u,t)}{\partial u}\mathrm{d}u\right| \leq \max_{u\in[n',n]}\left|\frac{\partial \mathbf{y}^{\mathsf{CL}}(u,t)}{\partial u}\right||n-n'|.$$

Because the analysis for $t=T$ (i.e., the last period) is different from the analysis for $t<T$, we analyze the two cases separately.

When $t=T$, we define the following partitions of the set $[0,\alpha]$:

$$S_1 = \left\{n \in [0,\alpha] : \frac{n}{\lambda(n,\alpha)} < \bar{y}\right\} \text{ and } S_2 = \left\{n \in [0,\alpha] : \frac{n}{\lambda(n,\alpha)} \geq \bar{y}\right\}.$$

When $t = T$, we have

$$\mathbf{y}^{\mathsf{CL}}(n,t) = \begin{cases} \frac{n}{\lambda(n,\alpha)} & \text{if } n \in S_1 \\ \bar{y} & \text{if } n \in S_2 \end{cases},$$

where $\bar{y}$ is defined in Lemma 1(ii). When $n \in S_1$, $\mathbf{y}^{\mathsf{CL}}(n,t)$ has bounded derivative w.r.t. n because of Lemma 1(iii). For $n \in S_2$, the function is constant, so the derivative is 0.

Now consider $t < T$. We will prove that the derivative of $\mathbf{y}^{\mathsf{CL}}(n,t)$ w.r.t. n is bounded for $n \in [0,\alpha]$. By definition, $\mathbf{y}^{\mathsf{CL}}(n,t) = y_0^{\mathsf{D}}(n, T-t+1)$ where

$$y_0^{\mathsf{D}}(n, T-t+1) = \arg\max_{y \le \frac{n}{\lambda(n,\alpha)}} R^{\mathsf{D},y}(n, T-t+1),$$

where $R^{\mathsf{D},y}(n,T') = x^{-1}(y)\lambda(n,\alpha)y + V^{\mathsf{D}}(T'; n - \lambda(n,\alpha)y, \alpha)$ was defined in (32).

By Claim 2, $R^{\mathsf{D},y}(n, T-t+1)$ is strictly concave in y for a given $(n, \alpha, T-t+1)$. Let $\bar{y}_{t,n}$ to be the value that satisfies

$$\frac{\partial}{\partial y} R^{\mathsf{D},y}(n, T-t+1)\Big|_{y=\bar{y}_{t,n}} = \lambda(n,\alpha)\frac{\partial}{\partial y}\left(x^{-1}(y)y\right)\Big|_{y=\bar{y}_{t,n}} - \lambda(n,\alpha)\frac{\partial V^{\mathsf{D}}(T-t; n', \alpha)}{\partial n'}\Big|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}} = 0,$$

so

$$\frac{\partial}{\partial y}\left(x^{-1}(y)y\right)\Big|_{y=\bar{y}_{t,n}} = \frac{\partial V^{\mathsf{D}}(T-t; n', \alpha)}{\partial n'}\Big|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}}. \tag{59}$$

Then, by defining

$$S_1' = \left\{ n \in [0,\alpha] : \frac{n}{\lambda(n,\alpha)} < \bar{y}_{t,n} \right\} \text{ and } S_2' = \left\{ n \in [0,\alpha] : \frac{n}{\lambda(n,\alpha)} \ge \bar{y}_{t,n} \right\},$$

we know

$$\mathbf{y}^{\mathsf{CL}}(n,t) = \begin{cases} \frac{n}{\lambda(n,\alpha)} & \text{if } n \in S_1' \\ \bar{y}_{t,n} & \text{if } n \in S_2'. \end{cases}$$

From Lemma 1(iii), the derivative of $\mathbf{y}^{\mathsf{CL}}(n,t)$ w.r.t. n is bounded when $n \in S_1'$. When $n \in S_2'$, the derivative of $\mathbf{y}^{\mathsf{CL}}(n,t) = \bar{y}_{t,n}$ w.r.t. n. We can now differentiate (59) with respect to n through chain rule. We let $\lambda_1(n,\alpha)$ denote the first-order partial derivative of $\lambda(n,\alpha)$ w.r.t. n. Specifically, we have

$$\frac{\partial \bar{y}_{t,n}}{\partial n}\left(x^{-1}(y)y\right)''\Big|_{y=\bar{y}_{t,n}} = \left(1 - \lambda_1(n,\alpha)\bar{y}_{t,n} - \lambda(n,\alpha)\frac{\partial \bar{y}_{t,n}}{\partial n}\right)\frac{\partial^2 V^{\mathsf{D}}(T-t; n', \alpha)}{\partial n'^2}\Big|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}}. \tag{60}$$

Rearranging terms in (60) yields the following relationship:

$$\left|\frac{\partial \bar{y}_{t,n}}{\partial n}\right| = \left| \frac{(1 - \lambda_1(n,\alpha)\bar{y}_{t,n})\frac{\partial^2 V^{\mathsf{D}}(T-t; n', \alpha)}{\partial n'^2}\Big|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}}}{\left(x^{-1}(y)y\right)''\Big|_{y=\bar{y}_{t,n}} + \lambda(n,\alpha)\frac{\partial^2 V^{\mathsf{D}}(T-t; n', \alpha)}{\partial n'^2}\Big|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}}} \right|. \tag{61}$$

The term on the RHS of (61) is bounded (i.e., the denominator is nonzero) because $r''(y) < 0$ is defined for $y \in [0,1]$ according to Lemma 1(i), $\partial^2 V^{\mathsf{D}}(T-t; n', \alpha)/\partial n'^2 < 0$ is defined for $n' \in [0,1]$ (Proposition 2(ii)), and $\lambda(n,\alpha)$ is continuous differentiable for $n \in [0,\alpha]$ and finite $\alpha \ge 0$. This concludes our proof. $\square$

C.12 Lemma 7 and proof

Before stating the lemma, we begin with introducing new notation.

For a given m, we define the stochastic sequence of inventory levels under the closed-loop policy as $\hat{N}^m = (\hat{N}_0^m, \hat{N}_1^m \ldots, \hat{N}_T^m)$, where $\hat{N}_0^m = \alpha m$. Recall that $\mathbf{y}^{\mathsf{CL}}$ sets the price in period t by optimizing the deterministic problem $(\mathbf{D}_m)$ on a rolling horizon, by replacing T with $T-t$ and setting $u = \hat{N}_{t-1}^m$. (As we discussed in Section 4.2, $(\mathbf{D}_m)$ is the scaled version of $(\mathbf{D})$. Hence, by the inventory constraint $(\mathbf{D}\text{b})$, the period t conditional expected demand under policy $\mathsf{CE\text{-}CL}$ would never exceed N_{t-1}^m.)

Lemma 7 (Convergence of remaining inventory and SIS)**.** If $n^{\mathsf{D}} = (n_1^{\mathsf{D}}, \ldots, n_T^{\mathsf{D}})$ is the solution to $(\mathbf{D})$ when $u = \alpha$, then the following hold:

$$\mathbb{E}\left|\frac{\hat{N}_t^m}{m} - n_t^{\mathsf{D}}\right| = \mathcal{O}\left(1/\sqrt{m}\right), \quad \text{for all } t = 1, \ldots, T \tag{62}$$

$$\mathbb{E}\left|\lambda\left(\frac{\hat{N}_t^m}{m}, \alpha\right) - \lambda\left(n_t^{\mathsf{D}}, \alpha\right)\right| = \mathcal{O}\left(1/\sqrt{m}\right), \quad \text{for all } t = 1, \ldots, T \tag{63}$$

Proof. The proof is analogous to that of Lemma 2 in Appendix C.8. We start by defining the sequence of random variables $(\hat{I}_0^m, \hat{I}_1^m, \ldots, \hat{I}_T^m)$, where $\hat{I}_t^m = \hat{N}_t^m/m$ is the normalized remaining inventory at time t under the closed-loop policy $\mathbf{y}^{\mathsf{CL}}$ when the initial inventory and the expected demand are scaled by m. Note that $\hat{I}_0^m = \alpha$.

We will prove the lemma by induction. The base case is $t = 0$, where we note that $\hat{I}_0^m = n_0^{\mathsf{D}} = \alpha$, and hence $\lambda(\hat{I}_0^m, \alpha) = \lambda(n_0^{\mathsf{D}}, \alpha) = \lambda(\alpha, \alpha)$. Therefore, (62) and (63) hold for $t = 0$. For the induction step, we assume that (62) and (63) hold for $t-1$. Specifically,

$$\mathbb{E}\left|\hat{I}_{t-1}^m - n_{t-1}^{\mathsf{D}}\right| = \mathcal{O}\left(1/\sqrt{m}\right), \tag{64}$$

$$\mathbb{E}\left|\lambda\left(\hat{I}_{t-1}^m, \alpha\right) - \lambda\left(n_{t-1}^{\mathsf{D}}, \alpha\right)\right| = \mathcal{O}\left(1/\sqrt{m}\right), \tag{65}$$

We will show these properties (62),(63) hold for t.

To prove (62) for t, notice that by adding and subtracting $\mathbb{E}(\hat{I}_t^m)$,

$$\mathbb{E}\left|\hat{I}_t^m - n_t^{\mathsf{D}}\right| = \mathbb{E}\left|\hat{I}_t^m - \mathbb{E}\left(\hat{I}_t^m\right) + \mathbb{E}\left(\hat{I}_t^m\right) - n_t^{\mathsf{D}}\right| \le \mathbb{E}\left|\hat{I}_t^m - \mathbb{E}\left(\hat{I}_t^m\right)\right| + \left|\mathbb{E}\left(\hat{I}_t^m\right) - n_t^{\mathsf{D}}\right|. \tag{66}$$

We will show that both terms in the right side of (66) are $\mathcal{O}(1/\sqrt{m})$.

Following the similar argument from the proof of Lemma 2 in Appendix C.8 until the end of proof of Claim 4, we have the first term on the RHS of (66) is $\mathcal{O}(1/\sqrt{m})$. For the second term in (66), we want to bound the difference between $\mathbb{E}(\hat{I}_t^m)$ and n_t^{D}. From the definition of $\hat{I}_t^m$, we know

$$\mathbb{E}(\hat{I}_t^m \mid \mathcal{F}_{t-1}) = \mathbb{E}\left(\left[\hat{I}_{t-1}^m - \frac{D_t^m}{m}\right]^+ \mid \mathcal{F}_{t-1}\right) = \frac{1}{m}\mathbb{E}\left(\left[\hat{N}_{t-1}^m - D_t^m\right]^+ \mid \mathcal{F}_{t-1}\right).$$

Note $\mathbb{E}(D_t^m \mid \mathcal{F}_{t-1}) = \lambda^m\left(\hat{N}_{t-1}^m, \alpha m\right)\mathbf{y}^{\mathsf{CL}}\left(\hat{N}_{t-1}^m, t\right)$ and, by Assumption 3, we also have a bound on the variance $\mathrm{Var}(D_t^m \mid \mathcal{F}_{t-1}) \le \sigma\lambda^m(\hat{N}_{t-1}^m, \alpha m)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)$. Therefore since $\hat{N}_{t-1}^m$ is not random when conditioning on the filtration $\mathcal{F}_{t-1}$, and using the Scarf bound and (15), we have

$$\mathbb{E}(\hat{I}_t^m \mid \mathcal{F}_{t-1})$$
$$\le \frac{1}{2}\left(\sqrt{\frac{\sigma\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)}{m} + \left(\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) - \hat{I}_{t-1}^m\right)^2} - \left(\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) - \hat{I}_{t-1}^m\right)\right)$$

$$\leq \frac{1}{2}\left(\sqrt{\frac{\sigma\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)}{m}} + \left|\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) - \hat{I}^m_{t-1}\right| - \left(\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) - \hat{I}^m_{t-1}\right)\right)$$

$$\leq \hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) + \frac{1}{2}\sqrt{\frac{\sigma\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)}{m}}.$$

The last inequality comes from the fact that given inventory level $\hat{N}^m_{t-1}$ at time t, the next price chosen by policy $\mathbf{y}^{\mathsf{CL}}$ always satisfies $\hat{N}^m_{t-1} - \lambda(\hat{N}^m_{t-1},\alpha m)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) \geq 0$ since it resolves (**D**) with updated inventory level $u = \hat{N}^m_{t-1}$ which has a constraint (**D**b) that the total expected demand cannot exceed inventory $\hat{N}^m_{t-1}$. Therefore, we have

$$\mathbb{E}(\hat{I}^m_t \mid \mathcal{F}_{t-1}) \leq \hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) + \Theta(m^{-\frac{1}{2}}). \tag{67}$$

Taking the expectation on both sides conditioning on $\mathcal{F}_0$, we have the upper bound

$$\mathbb{E}(\hat{I}^m_t) \leq \mathbb{E}\left(\hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right) + \Theta(1/\sqrt{m})$$

We also have a lower bound from the following arguments:

$$\begin{aligned}\mathbb{E}(\hat{I}^m_t) &= \mathbb{E}\left(\left(\hat{I}^m_{t-1} - \frac{D^m_t}{m}\right)^+\right)\\ &\geq \mathbb{E}\left(\hat{I}^m_{t-1} - \frac{D^m_t}{m}\right) = \mathbb{E}\left(\mathbb{E}\left(\hat{I}^m_{t-1} - \frac{D^m_t}{m} \mid \mathcal{F}_{t-1}\right)\right) = \mathbb{E}\left(\hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right),\end{aligned}$$

where the last relationship uses (15). Hence,

$$0 \leq \mathbb{E}\left(\hat{I}^m_t - \hat{I}^m_{t-1} + \lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right) \leq \Theta(1/\sqrt{m}). \tag{68}$$

This implies that

$$\begin{aligned}\left|\mathbb{E}(\hat{I}^m_t) - n^{\mathsf{D}}_t\right| &= \left|\mathbb{E}(\hat{I}^m_t) - n^{\mathsf{D}}_{t-1} + \lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right|\\ &\leq \left|\mathbb{E}\left(\hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right) - n^{\mathsf{D}}_{t-1} + \lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right| + \Theta(1/\sqrt{m}) \qquad (69)\\ &\leq \mathbb{E}\left|\hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) - n^{\mathsf{D}}_{t-1} + \lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right| + \Theta(1/\sqrt{m}) \qquad (70)\\ &\leq \mathbb{E}|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}| + \mathbb{E}|\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) - \lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t| + \Theta(1/\sqrt{m}) \qquad (71)\\ &\leq \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| + \underbrace{\mathbb{E}\left|\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) - \lambda(n^{\mathsf{D}}_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right|}_{(*)}\\ &\quad + \underbrace{\mathbb{E}\left|\lambda(n^{\mathsf{D}}_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) - \lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right|}_{(**)} + \Theta(1/\sqrt{m}), \qquad (72)\end{aligned}$$

where (69) follows from (68), (70) is from Jensen's inequality, (71) is from triangle inequality and monotonicity of expectation, (72) is derived by subtracting and adding $\lambda(n^{\mathsf{D}}_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)$ and using the triangle inequality.

To analyze the bound for $(*)$, we know λ is Lipschitz continuous. This is because λ is continuously differentiable in its two variables (Assumption 2(vi)), so there exists a C_λ such that $|\lambda(n,\alpha) - \lambda(n',\alpha)| \leq$

$C_\lambda|n-n'|$ for all n, n', and fixed α. Also, we know $\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) \le 1$ by Assumption 2(i). Therefore,

$$(*) \le 1 \cdot C_\lambda \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right|$$

To analyze the bound for $(**)$, we know from Lemma 4 that $\mathbf{y}^{\mathsf{CL}}(n,t)$ is Lipschitz continuous in n with some Lipschitz constant C_y. Furthermore, observe that $\mathbf{y}^{\mathsf{CL}}(mn^{\mathsf{D}}_t,t) = y^{\mathsf{D}}_t$. Another important property of $\mathbf{y}^{\mathsf{CL}}$ we need is that $\mathbf{y}^{\mathsf{CL}}(mn,t;m\alpha)$ under initial inventory is $m\alpha$ is the same as $\mathbf{y}^{\mathsf{CL}}(n,t;\alpha)$ under initial inventory is α. This is because $\mathbf{y}^{\mathsf{CL}}$ solves optimization model (**D**) where the optimal intensity is invariant under scaling since, for any $n \in [0,\alpha]$, $\lambda(mn,m\alpha) = m\lambda(n,\alpha)$ due to (15). Therefore,

$$\begin{aligned}(**) &= \mathbb{E}\left|\lambda(n^{\mathsf{D}}_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t;m\alpha) - \lambda(n^{\mathsf{D}}_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(mn^{\mathsf{D}}_t,t;m\alpha)\right| \\ &= \mathbb{E}\left|\lambda(n^{\mathsf{D}}_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{I}^m_{t-1},t;\alpha) - \lambda(n^{\mathsf{D}}_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(n^{\mathsf{D}}_t,t;\alpha)\right| \\ &\le \bar{\lambda} C_y \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| \end{aligned} \tag{73}$$

where the inequality is due to the Lipschitz continuity of $\mathbf{y}^{\mathsf{CL}}(n,t)$ in n, and because λ is upper bounded by $\bar{\lambda}$ according to Assumption 2(v). Therefore, we conclude

$$\begin{aligned}\left|\mathbb{E}(\hat{I}^m_t) - n^{\mathsf{D}}_t\right| &\le \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| + 1 \cdot C_\lambda \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| + \bar{\lambda} C_y \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| + \Theta(1/\sqrt{m}), \\ &= \mathcal{O}\left(1/\sqrt{m}\right), \end{aligned} \tag{74}$$

where (74) comes from the inductive hypothesis (64).

Therefore, we can conclude that the RHS two terms of (66) are both bounded by $\mathcal{O}(1/\sqrt{m})$, thus giving us (62) for all t by induction. For a given t, (63) follows by the Lipschitz continuity of λ and (62):

$$\mathbb{E}\left|\lambda(\hat{I}^m_t,\alpha) - \lambda(n^{\mathsf{D}}_t,\alpha)\right| \le C_\lambda \mathbb{E}\left|\hat{I}^m_t - n^{\mathsf{D}}_t\right| = \mathcal{O}\left(1/\sqrt{m}\right).$$

This concludes the proof. □

C.13 Lemma 8 and proof

An important implication of Lemma 7 is that the intensity policy $\mathbf{y}^{\mathsf{CL}}$ converges to the deterministic sequence $\mathbf{y}^{\mathsf{D}}$ since, with Lemma 4, we know that $\mathbf{y}^{\mathsf{CL}}$ is Lipschitz continuous. These properties allow us to show that the *uncensored* expected revenue under $\mathbf{y}^{\mathsf{CL}}$ has a gap from $V^{\mathsf{D}}(m,T)$ that is $\mathcal{O}(\sqrt{m})$. This is formalized in the lemma below.

Lemma 8 (Convergence of uncensored revenue)**.**

$$\left|\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right)\lambda^m(\hat{N}^m_{t-1},\alpha m)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right) - V^{\mathsf{D}}(m,T)\right| = \mathcal{O}\left(\sqrt{m}\right). \tag{75}$$

Proof. By definition of V^{D} and from property (15) of λ^m, $V^{\mathsf{D}}(m,T) = \sum_{t=1}^{T} x^{-1}(y^{\mathsf{D}}_t) m\lambda(n^{\mathsf{D}}_{t-1},\alpha) y^{\mathsf{D}}_t$. Hence, defining $\hat{I}^m_{t-1} = \hat{N}^m_{t-1}/m$, we can write the LHS of (75) as

$$\begin{aligned} &m\left|\mathbb{E}\left[\sum_{t=1}^{T}\left(x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) - x^{-1}(y^{\mathsf{D}}_t)\lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right)\right]\right| \\ &\le m\mathbb{E}\left|\sum_{t=1}^{T}\left(x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}^m_{t-1},t) - x^{-1}(y^{\mathsf{D}}_t)\lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right)\right| \end{aligned}$$

$$\leq m \sum_{t=1}^{T} \mathbb{E} \left| x^{-1} \left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) \right) \lambda(\hat{I}_{t-1}^m, \alpha) \mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) - x^{-1}(y_t^{\mathsf{D}}) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}} \right|. \tag{76}$$

Here, the first inequality comes from $|\mathbb{E}X| \leq \mathbb{E}|X|$ as a result of Jensen's inequality. The second inequality comes from triangle inequality and linearity of expectation. To prove the proposition, since T is a finite number, it is sufficient to show each term inside the summation of (76) is $\mathcal{O}\left(1/\sqrt{m}\right)$.

Note that for any t,

$$\begin{aligned}
&\mathbb{E} \left| x^{-1} \left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) \right) \lambda(\hat{I}_{t-1}^m, \alpha) \mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) - x^{-1}(y_t^{\mathsf{D}}) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}} \right| \\
&= \mathbb{E} \left| r \left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) \right) \lambda(\hat{I}_{t-1}^m, \alpha) - r(y_t^{\mathsf{D}}) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) \right|,
\end{aligned} \tag{77}$$

where $r(y) = x^{-1}(y)y$ is the per-period revenue rate. Our goal is to show that (77) is $\mathcal{O}(1/\sqrt{m})$.

We first prove the Lipschitz continuity of the function $r(y)$. From Lemma 1(i), $r(y)$ is concave in y and is continuously differentiable for $y \in [0,1]$. Therefore, there exists C_r such that

$$|r(y) - r(y')| \leq C_r |y - y'|. \tag{78}$$

Additionally, $r(y) \leq \bar{f} = r(\bar{y})$ where $\bar{y}$ is defined in Lemma 1(ii). Hence, if we subtract and add the term $r(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)) \lambda(n_{t-1}^{\mathsf{D}}, \alpha)$ inside the absolute value in (77), by triangle inequality, (77) is upper bounded by

$$\begin{aligned}
&\mathbb{E} \left| r \left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) \right) \lambda(\hat{I}_{t-1}^m, \alpha) - r \left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) \right) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) \right| \\
&\quad + \mathbb{E} \left| r \left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) \right) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) - r(y_t^{\mathsf{D}}) \lambda(n_{t-1}^{\mathsf{D}}, \alpha) \right| \\
&\leq \bar{f} \, \mathbb{E} \left| \lambda(\hat{I}_{t-1}^m, \alpha) - \lambda(n_{t-1}^{\mathsf{D}}, \alpha) \right| + \bar{\lambda} C_r \, \mathbb{E} \left| \mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) - y_t^{\mathsf{D}} \right|,
\end{aligned} \tag{79}$$

where the second term of (79) comes from (78) and Assumption 2(v). Hence, it suffices to show the two terms in (79) are bounded by $\mathcal{O}(1/\sqrt{m})$. This is true because, from (63) of Lemma 7, for any t,

$$\mathbb{E} \left| \lambda(\hat{I}_{t-1}^m, \alpha) - \lambda(n_{t-1}^{\mathsf{D}}, \alpha) \right| = \mathcal{O}\left(1/\sqrt{m}\right).$$

Moreover, by definition, $\mathbf{y}^{\mathsf{CL}}$ results from re-optimizing the deterministic equivalent at each time period, hence we have that $\mathbf{y}^{\mathsf{CL}}(mn_{t-1}^{\mathsf{D}}, t) = y_t^{\mathsf{D}}$. We use the property of $\mathbf{y}^{\mathsf{CL}}$ that $\mathbf{y}^{\mathsf{CL}}(mn, t; m\alpha)$ under initial inventory is $m\alpha$ is the same as $\mathbf{y}^{\mathsf{CL}}(n, t; \alpha)$ under initial inventory is α. This is because $\mathbf{y}^{\mathsf{CL}}$ solves optimization model (**D**) where the optimal deterministic intensity solution is invariant under scaling since (15) implies that, for any $n \in [0, \alpha]$, $\lambda^m(mn, m\alpha) = m\lambda(n, \alpha)$. Therefore,

$$\begin{aligned}
\mathbb{E} \left| \mathbf{y}^{\mathsf{CL}} \left(\hat{N}_{t-1}^m, t \right) - y_t^{\mathsf{D}} \right| &= \mathbb{E} \left| \mathbf{y}^{\mathsf{CL}} \left(\hat{N}_{t-1}^m, t; m\alpha \right) - \mathbf{y}^{\mathsf{CL}}(mn_{t-1}^{\mathsf{D}}, t; m\alpha) \right| \\
&= \mathbb{E} \left| \mathbf{y}^{\mathsf{CL}} \left(\hat{I}_{t-1}^m, t; \alpha \right) - \mathbf{y}^{\mathsf{CL}} \left(n_{t-1}^{\mathsf{D}}, t; \alpha \right) \right| \\
&\leq C_y \mathbb{E} \left| \hat{I}_{t-1}^m - n_{t-1}^{\mathsf{D}} \right| = \mathcal{O}\left(1/\sqrt{m}\right),
\end{aligned}$$

where the inequality is from Lemma 4, and the last equality is from (62) of Lemma 7. This concludes the proof. □

C.14 Proof of Theorem 3

Proof. First, we note that $V^*(m,T)$ is greater or equal to the revenue from a single-price policy, and so is strictly positive. To prove the theorem, it is sufficient to show that

$$1 - \frac{V^{\mathsf{CL}}(m,T)}{V^{\mathsf{D}}(m,T)} \le 1 - (1-k)\left(1 - \frac{C}{2}\sqrt{\frac{\sigma}{m}}\right), \tag{80}$$

where $k = \Theta(1/\sqrt{m})$ and C is some constant that is independent of m.

Recall $\hat{N}^m = (\hat{N}_0^m, \ldots, \hat{N}_T^m)$ is the stochastic sequence of remaining inventories under $\mathbf{y}^{\mathsf{CL}}$, where initial inventory is $\hat{N}_0^m = \alpha m$. Then from (6), we have

$$V^{\mathsf{CL}}(m,T) = \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{E}\left[x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)\right)(D_t^m - [D_t^m - \hat{N}_{t-1}^m]^+) \mid \mathcal{F}_{t-1}\right]\right]. \tag{81}$$

We next define the random variable $\hat{I}_t^m \triangleq \hat{N}_t^m/m$ for all t, where $\hat{I}_0^m = \alpha$. Note that $\hat{N}_{t-1}^m, \hat{I}_{t-1}^m$ are not random when conditioning on the filtration $\mathcal{F}_{t-1}$. Further, $\mathbb{E}(D_t^m \mid \mathcal{F}_{t-1}) = m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)$ and, by Assumption 3, $\mathrm{Var}(D_t^m \mid \mathcal{F}_{t-1}) \le \sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)$. Therefore, by the Scarf bound and from (15) we have

$$\begin{aligned}
&\mathbb{E}\left[[D_t^m - \hat{N}_{t-1}^m]^+ \mid \mathcal{F}_{t-1}\right] \\
&\le \frac{1}{2}\left(\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) + \left(\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)\right)^2} - \left(\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)\right)\right)
\end{aligned}$$

If we multiply the numerator and denominator of the right-hand side by the same term

$$\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) + \left(\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)\right)^2} + \left(\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)\right),$$

then we have the following:

$$\begin{aligned}
&\mathbb{E}\left[[D_t^m - \hat{N}_{t-1}^m]^+ \mid \mathcal{F}_{t-1}\right] \\
&\le \frac{1}{2}\frac{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)}{\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) + (\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t))^2} + \left(\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)\right)} \\
&\le \frac{1}{2}\frac{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)}{\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) + (\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t))^2}} \\
&\le \frac{1}{2}\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)}.
\end{aligned} \tag{82}$$

The second inequality holds because, conditional on $\mathcal{F}_{t-1}$, $\hat{N}_{t-1}^m - \lambda(\hat{N}_{t-1}^m, \alpha m)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) \ge 0$. This is because, at time t, the closed-loop policy solves the deterministic problem (**D**) with parameter $u = \hat{N}_{t-1}^m$ and initial inventory αm, which has a constraint that the expected demand cannot exceed u.

Therefore, plugging (82) into the RHS of (81), we observe that

$$V^{\mathsf{CL}}(m,T) \ge \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)\right)\left(m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t) - \frac{1}{2}\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^m, t)}\right)\right]$$

$$= \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right) m\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right]$$

$$- \frac{1}{2}\sqrt{\sigma m}\, \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right)\sqrt{\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)}\right]$$

$$= \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right) m\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right]$$

$$\times \left(1 - \frac{1}{2}\sqrt{\frac{\sigma}{m}}\underbrace{\frac{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right)\sqrt{\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)}\right]}{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right)\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right]}}_{(**)}\right). \tag{83}$$

We get the first equality by multiplying x^{-1} term inside. The second equality comes from pulling out the first expectation term.

We first derive a lower bound for the first term in (83). Note that $\mathbf{y}^{\mathsf{CL}}$ does not scale with m since it is constructed from the intensity solution of (**D**), which is scale-invariant due to property (15) of λ. From Lemma 8, we know that the difference between the first term in (83) and $V^{\mathsf{D}}(m,T)$ scales in $\mathcal{O}(\sqrt{m})$. This is slower than the speed of scaling $\Theta(m)$ of $V^{\mathsf{D}}(m,T)$. Hence,

$$\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right) m\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right] \geq V^{\mathsf{D}}(m,T)(1-k) \tag{84}$$

where $k = \Theta(1/\sqrt{m})$.

Next, we want to derive an upper bound for the term $(**)$, which results in a lower bound for the second term in (83). From Cauchy-Swartz inequality, the numerator of $(**)$ is bounded above by

$$\mathbb{E}\left[\sqrt{\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right)\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)}\sqrt{\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right)}\right]$$

$$\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right)\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)}\right]\sqrt{Tx^{-1}(0)}$$

$$\leq \sqrt{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right)\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right]}\sqrt{Tx^{-1}(0)},$$

where the first inequality comes from Assumption 2(ii), and the last inequality comes from Jensen's inequality and the fact that $\sqrt{z}$ is a concave function. Hence,

$$(**) \leq \sqrt{\frac{Tx^{-1}(0)}{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right)\lambda(\hat{I}_{t-1}^{m},\alpha)\mathbf{y}^{\mathsf{CL}}(\hat{N}_{t-1}^{m},t)\right]}} \leq \sqrt{\frac{Tx^{-1}(0)}{V^{\mathsf{D}}(T)(1-k)}},$$

where the last inequality comes from (84).

Since $\Theta(1/\sqrt{m})$ decreases as m grows, we know there exists some constant $\Theta(1)$ unaffected by m such that $\Theta(1/\sqrt{m}) \leq \Theta(1)$. Therefore, we know

$$\sqrt{\frac{1}{1-k}} = \sqrt{\frac{1}{1-\Theta(1/\sqrt{m})}} \leq \sqrt{\frac{1}{1-\Theta(1)}} = \Theta(1).$$

Hence, we have that

$$(**) \leq \sqrt{\frac{Tx^{-1}(0)}{V^{\mathsf{D}}(T)}}\Theta(1) \triangleq C. \tag{85}$$

Finally, we take (84) and (85) into (83), resulting in

$$V^{\mathsf{CL}}(m,T) \geq V^{\mathsf{D}}(m,T)(1-k)\left(1-\frac{1}{2}\sqrt{\frac{\sigma}{m}}C\right).$$

This completes the proof. □

Appendix D Section 5 proofs

D.1 Proof of Theorem 4

Proof. We denote as $(\alpha^*, \mathbf{y}^*)$ the optimal inventory and pricing policy of the stochastic problem ($\mathbf{P}'$) for some demand process that satisfies Assumptions 1 to 3. Because $Q^{\mathsf{CE}}(m,T) = V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{CE}}}(m,T) - c\alpha^{\mathsf{CE}}m$, we first analyze the bound for $V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{CE}}}(m,T)$ and then get $Q^{\mathsf{CE}}(m,T)$ by subtracting $c\alpha^{\mathsf{CE}}m$.

Let $(N_0^m, N_1^m, \ldots, N_T^m)$ be the sequence of stochastic remaining inventories under the joint initial inventory and pricing policy $(m\alpha^{\mathsf{CE}}, \mathbf{y}^{\mathsf{CE}})$. Define $I_t^m \triangleq N_t^m/m$. From (83) and (85), we know

$$V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{CE}}}(m,T) \geq \mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}\left(N_{t-1}^m,t\right)\right)m\lambda\left(I_{t-1}^m,\alpha^{\mathsf{CE}}\right)\mathbf{y}^{\mathsf{CE}}\left(N_{t-1}^m,t\right)\right)\left(1-\frac{1}{2}\sqrt{\frac{\sigma}{m}}C\right). \tag{86}$$

Note that Lemmas 3 and 8 implies that

$$\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}\left(N_{t-1}^m,t\right)\right)m\lambda\left(I_{t-1}^m,\alpha^{\mathsf{CE}}\right)\mathbf{y}^{\mathsf{CE}}\left(N_{t-1}^m,t\right)\right) \geq m\left(V^{\mathsf{D},\alpha^{\mathsf{CE}}}(T)-k\right), \tag{87}$$

where $k = \mathcal{O}\left(1/\sqrt{m}\right)$ and $k \geq 0$. Therefore, subtracting both sides of (86) by $c\alpha^{\mathsf{CE}}m$, and using (87), we have

$$\underbrace{V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{CE}}}(m,T) - c\alpha^{\mathsf{CE}}m}_{Q^{\mathsf{CE}}(m,T)} \geq m\left(V^{\mathsf{D},\alpha^{\mathsf{CE}}}(T)-k\right)\left(1-\frac{1}{2}\sqrt{\frac{\sigma}{m}}C\right) - c\alpha^{\mathsf{CE}}m. \tag{88}$$

Now, we analyze the RHS of (88) to connect it to $Q^*(m,T)$. Define $k_1 = \frac{1}{2}\sqrt{\frac{\sigma}{m}}C$ where C is defined in (85) with $\alpha = \alpha^{\mathsf{CE}}$.

Factoring out $m(1-k_1)$ in the RHS of (88) results in

$$m(1-k_1)\left(V^{\mathsf{D},\alpha^{\mathsf{CE}}}(T) - k - \frac{c\alpha^{\mathsf{CE}}}{1-k_1}\right)$$

$$
\begin{aligned}
&= m(1-k_1)\left(\underbrace{V^{\mathsf{D},\alpha^{\mathsf{CE}}}(T) - c\alpha^{\mathsf{CE}}}_{Q^{\mathsf{D},\alpha^{\mathsf{CE}}}(T)} - k + c\alpha^{\mathsf{CE}} - \frac{c\alpha^{\mathsf{CE}}}{1-k_1}\right) && \text{subtracting and adding } c\alpha^{\mathsf{CE}} \\
&\geq m(1-k_1)\left(\underbrace{V^{\mathsf{D},\alpha^*}(T) - c\alpha^*}_{Q^{\mathsf{D},\alpha^*}(T)} - k - c\alpha^{\mathsf{CE}}\frac{k_1}{1-k_1}\right) && \text{definition of } \alpha^{\mathsf{CE}} \text{ so } Q^{\mathsf{D},\alpha^{\mathsf{CE}}}(T) \geq Q^{\mathsf{D},\alpha^*}(T) \\
&= (1-k_1)\left(V^{\mathsf{D},\alpha^*}(m,T) - c\alpha^* m - mk - c\alpha^{\mathsf{CE}} m\frac{k_1}{1-k_1}\right) && \text{multiplying } m \text{ inside} \\
&\geq (1-k_1)\left(V^{\alpha^*,\mathbf{y}^*}(m,T) - c\alpha^* m - mk - c\alpha^{\mathsf{CE}} m\frac{k_1}{1-k_1}\right) && \text{from Proposition 5} \\
&= (1-k_1)\left(\underbrace{V^{\alpha^*,\mathbf{y}^*}(m,T) - c\alpha^* m}_{Q^*(m,T)} - (m + c\alpha^{\mathsf{CE}} m)k_2\right) && (89)
\end{aligned}
$$

with $k_2 = \Theta\left(1/\sqrt{m}\right)$ because

$$\frac{k_1}{1-k_1} = \Theta\left(\frac{1}{\sqrt{m}-1}\right).$$

Dividing (88) and the RHS of (89) by $Q^*(m,T) = V^{\alpha^*,\mathbf{y}^*}(m,T) - c\alpha^* m$ yields

$$\frac{Q^{\mathsf{CE}}(m,T)}{Q^*(m,T)} \geq (1-k_1)\left(1 - k_2 \cdot \frac{m + c\alpha^{\mathsf{CE}} m}{V^{\alpha^*,\mathbf{y}^*}(m,T) - c\alpha^* m}\right).$$

Hence, to prove (22), it suffices to show

$$\frac{m + c\alpha^{\mathsf{CE}} m}{V^{\alpha^*,\mathbf{y}^*}(m,T) - c\alpha^* m} = \mathcal{O}(1).$$

This is true because

$$\frac{m + c\alpha^{\mathsf{CE}} m}{V^{\alpha^*,\mathbf{y}^*}(m,T) - c\alpha^* m} = \frac{m\left(1 + c\alpha^{\mathsf{CE}}\right)}{m\left(V^{\mathsf{D}}(T) - \mathcal{O}(1/\sqrt{m}) - c\alpha^*\right)} = \Theta\left(\frac{1 + c\alpha^{\mathsf{CE}}}{V^{\mathsf{D}}(T) - c\alpha^*}\right),$$

which is constant in m. This concludes the proof. □

D.2 Proof of Proposition 6

Proof. Since it is not possible to characterize the exact revenue difference between the optimal and the fixed price policy, to prove Proposition 6, we utilize the bound established by V^{D}. To see this, an implication of our results in Section 3 is $0 \leq V^{\mathsf{D}}(m,T) - V^*(m,T) \leq \mathcal{O}(\sqrt{m})$ (Proposition 5, Theorem 3). In other words, $V^{\mathsf{D}}(m,T)$ is a good approximation of the optimal revenue in an asymptotic regime. Hence, if we are able to show for any $\alpha \geq 0$ that

$$V^{\mathsf{D},\alpha}(m,T) - V^{\mathsf{FP},\alpha}(m,T) = \Omega(m(T-1)), \tag{90}$$

then this establishes the first statement in Proposition 6. Note that this also proves the second statement because $Q^*(m,T) - Q^{\mathsf{FP}}(m,T) \geq V^{\mathsf{D},\alpha^{\mathsf{FP}}}(m,T) - V^{\mathsf{FP},\alpha^{\mathsf{FP}}}(m,T)$.

We need two key results to prove (90). The first key result in establishing (90) is to show that $V^{\mathsf{D},\alpha}(m,T) - V^{\mathsf{D}',\alpha}(m,T) = \Theta(m)$, where $V^{\mathsf{D}',\alpha}(m,T)$ is the deterministic revenue under the fixed price defined in (24) when the initial inventory is αm. This is formalized in the following lemma.

Lemma 9 (Revenue loss of the fixed price policy for deterministic problems)**.** When $T \geq 2$, for a fixed $\alpha \geq 0$, if

(i) $\frac{\partial}{\partial y} V^{\mathsf{D}}(T-1; \alpha - \lambda(\alpha,\alpha)y, \alpha)\Big|_{y=\bar{y}} \neq 0$, and

(ii) $\alpha \geq \sum_{t=1}^{T} \lambda(n_{t-1}^{\bar{y}}, \alpha)\bar{y}$,

then $V^{\mathsf{D},\alpha}(m,T) - V^{\mathsf{D}',\alpha}(m,T) = \Theta(m(T-1))$.

Proof. Consider an arbitrary $\alpha \geq 0$ satisfying the conditions of the lemma. Recall the definition $R^{\mathsf{D}}(u,T)$ in (31), where $V^{\mathsf{D}}(T) = R^{\mathsf{D}}(\alpha,T)$.

Due to condition (ii) of the lemma and from (23), we have that $y^{\mathsf{FP}} = \bar{y}$. Define the recursive equations

$$R^{\mathsf{D}'}(u,T) = x^{-1}(\bar{y})\lambda(u,\alpha)\bar{y} + R^{\mathsf{D}'}(\alpha - \lambda(u,\alpha)\bar{y}, T-1),$$

where $R^{\mathsf{D}'}(u,0) = 0$ for all $u \in [0,\alpha]$. Note that $V^{\mathsf{D}'}(T) = R^{\mathsf{D}'}(\alpha,T)$.

We next define

$$R^{\mathsf{D},y}(u,T) \triangleq x^{-1}(y)\lambda(u,\alpha)y + R^{\mathsf{D}}(\alpha - \lambda(u,\alpha)y, T-1) \text{ and}$$
$$R^{\mathsf{D}',y}(u,T) \triangleq x^{-1}(y)\lambda(u,\alpha)y + R^{\mathsf{D}'}(\alpha - \lambda(u,\alpha)y, T-1),$$

where $R^{\mathsf{D}}(u,T)$ is defined in (31). Note that $R^{\mathsf{D},y}(u,T)$ is the objective in (31). From the definition of y_1^{D}, when $u=\alpha$, $R^{\mathsf{D},y}(\alpha,T)$ achieves its maximum value $V^{\mathsf{D}}(T)$ when $y = y_1^{\mathsf{D}}$. We observe that

$$R^{\mathsf{D}}(\alpha,T) - R^{\mathsf{D}'}(\alpha,T) = \underbrace{R^{\mathsf{D},y_1^{\mathsf{D}}}(\alpha,T) - R^{\mathsf{D},\bar{y}}(\alpha,T)}_{(a)} + \underbrace{R^{\mathsf{D},\bar{y}}(\alpha,T) - R^{\mathsf{D}',\bar{y}}(\alpha,T)}_{(b)}. \tag{91}$$

We first note that in (91), $(b) \geq 0$ because

$$(b) = R^{\mathsf{D}}(\alpha - \lambda(\alpha,\alpha)\bar{y}, T-1) - R^{\mathsf{D}'}(\alpha - \lambda(\alpha,\alpha)\bar{y}, T-1) \geq 0$$

because $R^{\mathsf{D}}(\cdot,\cdot) = V^{\mathsf{D}}(\cdot,\cdot)$ defined in (**D**) is the optimal value, and $R^{\mathsf{D}'}(\cdot,\cdot)$ is the objective value of model (**D**) when $y_t = \bar{y}$ for all t (we can check that $\bar{y}$ is feasible to (**D**)).

For (a), because $R^{\mathsf{D},y}$ is strictly concave in y (Claim 2) and $y_1^{\mathsf{D}} > 0$ (Proposition 4), then we know

$$\frac{\partial R^{\mathsf{D},y}(\alpha,T)}{\partial y}\Big|_{y\,=\,y_1^{\mathsf{D}}} = \underbrace{\frac{\partial}{\partial y}x^{-1}(y)\lambda(\alpha,\alpha)y\Big|_{y\,=\,y_1^{\mathsf{D}}}}_{(c)} + \underbrace{\frac{\partial}{\partial y}R^{\mathsf{D}}(\alpha - \lambda(\alpha,\alpha)y, T-1)\Big|_{y\,=\,y_1^{\mathsf{D}}}}_{(d)} = 0. \tag{92}$$

Condition (i) of Lemma 9 states that $(d) \neq 0$ which, combined with (92), implies that $(c) \neq 0$. Because $\bar{y}$ is the unique value that can make $\frac{\partial}{\partial y}x^{-1}(y)\lambda(\alpha,\alpha)y$ equal to zero (Lemma 1(ii)), we conclude $y_1^{\mathsf{D}} \neq \bar{y}$. Therefore, by the mean value theorem, there exists a $y' \in \big(\min\{\bar{y}, y_1^{\mathsf{D}}\}, \max\{\bar{y}, y_1^{\mathsf{D}}\}\big)$ such that

$$(a) = R^{\mathsf{D},y_1^{\mathsf{D}}}(\alpha,T) - R^{\mathsf{D},\bar{y}}(\alpha,T) = \frac{\partial R^{\mathsf{D},y}(\alpha,T)}{\partial y}\Big|_{y=y'}(y_1^{\mathsf{D}} - \bar{y}). \tag{93}$$

Note that $(a) \geq 0$ because y_1^{D} is the maximizer of $R^{\mathsf{D},y}(\alpha,T)$. Note that the derivative term in (93) is nonzero because $y' \neq y_1^{\mathsf{D}}$ and y_1^{D} is the unique maximizer of $R^{\mathsf{D},y}(\alpha, T)$. Further, since $y_1^{\mathsf{D}} \neq \bar{y}$, we have that $(a) > 0$.

We then use induction to show that $R^{\mathsf{D}}(\alpha, T) - R^{\mathsf{D}'}(\alpha, T) = \Omega(T-1)$ for $T \geq 2$.

Base step: When $T = 2$, it is proved above because $(a) > 0$.

Inductive step: For induction assumption, we assume $R^{\mathsf{D}}(\alpha, T) - R^{\mathsf{D}'}(\alpha, T) = \Omega(T-1)$ holds for $T = t$. We will prove it holds for $T = t+1$.

From (91), we know

$$R^{\mathsf{D}}(\alpha, t+1) - R^{\mathsf{D}'}(\alpha, t+1) = (a) + (b) = (a) + \Omega(t-1) = \Omega(t).$$

This concludes the proof of Lemma 9. □

Condition (i) of Lemma 9 implies the myopic optimal intensity $\bar{y}$ is not the optimal first-period price the deterministic model $V^{\mathsf{D}}(T)$. Condition (ii) means that we have a sufficient amount of initial inventory if we use to set the price at $x^{-1}(\bar{y})$.

The second key piece is the following lemma, which can be established from results in Section 3, that the gap between the expected revenue $V^{\mathsf{FP},\alpha}(m, T)$ and the deterministic revenue $V^{\mathsf{D}',\alpha}(m, T)$ is $\mathcal{O}(\sqrt{m})$.

Lemma 10. For a fixed $\alpha \geq 0$,

$$V^{\mathsf{FP},\alpha}(m, T) \leq V^{\mathsf{D}',\alpha}(m, T) + \mathcal{O}(\sqrt{m}).$$

Proof. Given $\alpha \geq 0$, let $(N_0^m, N_1^m, \ldots, N_T^m)$ denote the sequence of stochastic remaining inventory under policy $\mathbf{y}^{\mathsf{FP}}$ with $N_0^m = \alpha m$. Define $I_t^m \triangleq N_t^m / m$.

First we notice that

$$V^{\alpha,\mathbf{y}^{\mathsf{FP}}}(m, T) \leq m\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}(y^{\mathsf{FP}})\lambda\left(I_{t-1}^m, \alpha\right) y^{\mathsf{FP}}\right) \tag{94}$$

because the RHS is the expected revenue under $\mathbf{y}^{\mathsf{FP}}$ without the inventory constraint.

According to Corollary 3 (see Appendix D.4), we know

$$mV^{\mathsf{FP}}(\alpha, T) - \mathcal{O}(\sqrt{m}) \leq m\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}(\mathbf{y}^{\mathsf{FP}})\lambda\left(I_{t-1}^m, \alpha\right)\mathbf{y}^{\mathsf{FP}}\right) \leq mV^{\mathsf{FP}}(\alpha, T) + \mathcal{O}(\sqrt{m}). \tag{95}$$

Plugging (95) into RHS of (94), we get

$$V^{\alpha,\mathbf{y}^{\mathsf{FP}}}(m, T) \leq mV^{\mathsf{FP},\alpha}(T) + \mathcal{O}(\sqrt{m}) = V^{\mathsf{FP},\alpha}(m, T) + \mathcal{O}(\sqrt{m}).$$

Hence, $V^{\mathsf{D}}(T) - V^{\mathsf{FP}}(T) > 0$. This implies that $V^{\mathsf{D}}(m, T) - V^{\mathsf{FP}}(m, T) = m\left(V^{\mathsf{D}}(T) - V^{\mathsf{FP}}(T)\right) = \Theta(m)$. This concludes our proof. □

Now, we are ready to finish the proof of Proposition 6. From the definition that $Q^*(m, T)$ is the optimal profit, we know $Q^*(m, T) \geq V^{*,\alpha^{\mathsf{FP}}}(m, T) - m\alpha^{\mathsf{FP}}c$. Then,

$$\begin{aligned} Q^*(m, T) - Q^{\mathsf{FP}}(m, T) &\geq \left(V^{*,\alpha^{\mathsf{FP}}}(m, T) - m\alpha^{\mathsf{FP}}c\right) - \left(V^{\mathsf{FP},\alpha^{\mathsf{FP}}}(m, T) - m\alpha^{\mathsf{FP}}c\right) \\ &= V^{*,\alpha^{\mathsf{FP}}}(m, T) - V^{\mathsf{FP},\alpha^{\mathsf{FP}}}(m, T). \end{aligned}$$

Hence, to prove the proposition, it suffices to show $V^{*,\alpha}(m,T) - V^{\mathsf{FP},\alpha}(m,T) = \Omega(m(T-1))$ for any fixed $\alpha \geq 0$.

We know that $V^{*,\alpha}(m,T)$ is bounded below by $V^{\mathsf{CL},\alpha}(m,T)$. Hence, by Theorem 6, we have that $V^{*,\alpha}(m,T) \geq V^{\mathsf{D},\alpha}(m,T) - \mathcal{O}(\sqrt{m})$. This and Lemma 10 result in

$$V^{*,\alpha}(m,T) - V^{\mathsf{FP},\alpha}(m,T) \geq V^{\mathsf{D},\alpha}(m,T) - \mathcal{O}(\sqrt{m}) - V^{\mathsf{D}',\alpha}(m,T) - \mathcal{O}(\sqrt{m}). \tag{96}$$

Moreover, according to Lemma 9, we know the RHS of (96) is $\Omega\left(m(T-1)\right)$. This concludes the proof of Proposition 6. □

D.3 Corollary 2 and proof

Corollary 2. Given $\alpha \in [0,1]$, let $(N_0^{\mathsf{FP}}, N_1^{\mathsf{FP}}, \ldots, N_T^{\mathsf{FP}})$ denote the sequence of stochastic remaining inventory under policy $\mathbf{y}^{\mathsf{FP}}$ with $N_0^{\mathsf{FP}} = \alpha m$. Define $n_t^{\mathsf{FP}} \triangleq N_t^{\mathsf{FP}}/m$. Let $n^{\mathsf{D},\mathsf{FP}} = (n_0^{\mathsf{D},\mathsf{FP}}, \ldots, n_T^{\mathsf{D},\mathsf{FP}})$ be the deterministic optimal solution of (**D**) when fixing $y = (y^{\mathsf{FP}}, \ldots, y^{\mathsf{FP}})$. Then the following hold:

$$\mathbb{E}\left|n_t^{\mathsf{FP}} - n_t^{\mathsf{D},\mathsf{FP}}\right| = \mathcal{O}\left(m^{-\frac{1}{2}}\right)$$

and

$$\mathbb{E}\left|\lambda\left(n_t^{\mathsf{FP}}, \alpha\right) - \lambda\left(n_t^{\mathsf{D},\mathsf{FP}}, \alpha\right)\right| = \mathcal{O}\left(m^{-\frac{1}{2}}\right).$$

Proof. The only difference between Corollary 2 and Lemma 7 is the gap between the stochastic intensity sequence and the deterministic intensity sequence. In Lemma 7 (using the notation in the proof of Lemma 7), we apply $\mathbf{y}^{\mathsf{CL}}$ to the stochastic problem and accordingly get n^{CL}; and we apply y^{D} to the deterministic problem and accordingly have n^{D}. However, in Corollary 2, we apply $(y^{\mathsf{FP}}, \ldots, y^{\mathsf{FP}})$ to the stochastic problem and accordingly get n^{FP}; and we apply the same $(y^{\mathsf{FP}}, \ldots, y^{\mathsf{FP}})$ to the deterministic problem and accordingly have $n^{\mathsf{D},\mathsf{FP}}$. As a result, the key difference between the proofs of Lemma 7 and Corollary 2 is the logic to have the same $(**)$ in (72) upper bounded by $(****)$ in (73). Note that the definition of $\mathbf{y}^{\mathsf{FP}}$ in (23) also guarantees that inventory constraint is satisfied in expectation, so the logic in the proof stays the same as Lemma 7.

In Lemma 7, (using the notation in the proof of Lemma 7) we have the gap between $\mathbf{y}^{\mathsf{CL}}$ and y^{D} is

$$\mathbb{E}\left|\mathbf{y}^{\mathsf{CL}}\left(n_{t-1}^{\mathsf{CL}}, t\right) - y_t^{\mathsf{D}}\right| \leq \bar{\lambda} C_y \mathbb{E}\left|n_{t-1}^{\mathsf{CL}} - n_{t-1}^{\mathsf{D}}\right| = \mathcal{O}\left(m^{-\frac{1}{2}}\right). \tag{97}$$

Note that (97) is the key to have $(**) \leq (****)$ in the proof of Lemma 7. To get (97), the crucial part is the Lipschitz continuity of policy y^{CL} proved in Lemma 4. Therefore, in Corollary 2, if we also have the gap between y sequences applied to the stochastic and deterministic problems is $\mathcal{O}(m^{-\frac{1}{2}})$, then we are done. In fact, for Corollary 2, we apply the same sequence $(y^{\mathsf{FP}}, \ldots, y^{\mathsf{FP}})$ to both stochastic and deterministic problems, so clearly

$$\mathbb{E}\left|\mathbf{y}^{\mathsf{FP}}\left(n_{t-1}^{\mathsf{FP}}, t\right) - y^{\mathsf{FP}}\right| = 0,$$

thus is $\mathcal{O}(m^{-\frac{1}{2}})$. Therefore, we get the same bound as (73) in the proof of Lemma 7. Then, Corollary 2 holds by applying the same logic as the proof of Lemma 7.

□

D.4 Corollary 3 and proof

Corollary 3. Given $\alpha \in [0,1]$, let $(N_0^{\mathsf{FP}}, N_1^{\mathsf{FP}}, \dots, N_T^{\mathsf{FP}})$ denote the sequence of stochastic remaining inventory under policy $\mathbf{y}^{\mathsf{FP}}$ with $N_0^{\mathsf{FP}} = \alpha m$. Define $n_t^{\mathsf{FP}} \triangleq N_t^{\mathsf{FP}}/m$. Then,

$$\left| \mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{FP}}(n_{t-1}^{\mathsf{FP}}, t)\right) \lambda\left(n_{t-1}^{\mathsf{FP}}, \alpha\right) \mathbf{y}^{\mathsf{FP}}(n_{t-1}^{\mathsf{FP}}, t) \right) - V^{\mathsf{FP}}(\alpha, T) \right| = \mathcal{O}\left(m^{-\frac{1}{2}}\right).$$

Proof. Similar to the proof of Corollary 2 (see Appendix D.3), the only difference between Corollary 3 and Lemma 8 is the gap between the stochastic intensity sequence and the deterministic intensity sequence. In Lemma 8 (using the notation in the proof of Lemma 8), we apply $\mathbf{y}^{\mathsf{CL}}$ to the stochastic problem and accordingly get n^{CL} and the expected revenue

$$\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CL}}\left(n_{t-1}^{\mathsf{CL}}, t\right)\right) \lambda\left(n_{t-1}^{\mathsf{CL}}, \alpha\right) \mathbf{y}^{\mathsf{CL}}\left(n_{t-1}^{\mathsf{CL}}, t\right) \right);$$

and we apply y^{D} to the deterministic problem (**D**) and accordingly have n^{D} and the deterministic revenue $V^{\mathsf{D}}(\alpha, T)$. However, in Corollary 2, we apply $(y^{\mathsf{FP}}, \dots, y^{\mathsf{FP}})$ to the stochastic problem and accordingly get n^{FP} and the expected revenue

$$\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}(y^{\mathsf{FP}}) \lambda\left(n_{t-1}^{\mathsf{FP}}, \alpha\right) y^{\mathsf{FP}} \right);$$

and we apply the same $(y^{\mathsf{FP}}, \dots, y^{\mathsf{FP}})$ to the deterministic problem (**D**) and accordingly have $n^{\mathsf{D,FP}}$ and the deterministic revenue $V^{\mathsf{FP}}(\alpha, \alpha, T)$.

The proof of Corollary 3 follows exactly the same logic of the proof of Lemma 8. Whenever we use Lemma 7 in the proof of Lemma 8, we replace these with Corollary 2. Whenever we use Lemma 4 to bound $\mathbb{E}\left|\mathbf{y}^{\mathsf{CL}}\left(n_{t-1}^{\mathsf{CL}}, t\right) - y_t^{\mathsf{D}}\right|$, we do not need them because we have zero gap between two sequences of y, that is $\mathbb{E}\left|\mathbf{y}^{\mathsf{FP}}\left(n_{t-1}^{\mathsf{FP}}, t\right) - y^{\mathsf{FP}}\right| = 0$. $\square$