Revenue Management in the New Age: Analysis and Learning with Dependency and Non-Stationarity

by

Mengzhenyu Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in The University of Michigan
2021

Doctoral Committee:

Professor Hyun-Soo Ahn, Co-Chair
Assistant Professor Joline Uichanco, Co-Chair
Assistant Professor Christopher Thomas Ryan
Associate Professor Eric Schwartz
Associate Professor Cong Shi

Mengzhenyu Zhang
zhenyuzh@umich.edu
ORCID iD: 0000-0002-6069-576X

TABLE OF CONTENTS

LIST OF FIGURES

<u>**Figure**</u>

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

In this dissertation, we focus on revenue management practices in the settings where the traditional assumptions are challenged because of the rise of social technologies in the new age. We develop theoretically sound and practical policies (pricing, inventory, and information signals) when the firm faces uncertainty, dependency and non-stationarity in market demand. We also provide effective insights into the firm's decision making.

The first part of the dissertation is motivated by the challenges of decision making for *a new product*. We study the interplay between pricing and learning for a monopolist whose objective is to maximize the expected revenue of a new product over a finite time horizon. We consider a setting where a firm can learn by observing sales data at (different) prices over time. Its customers are not forward-looking. To capture the adoption process with demand learning, we develop a continuous-time Markovian Bass model where the adoption rate depends on the selling price and on the past sales. We first show that the revenue loss from a wrong initial demand estimation can grow proportionally in time and can be significant. We then show that a firm can apply real-time demand data to update the demand parameters and optimize the price accordingly. We formulate the problem as a stochastic optimal control problem where the demand parameters are updated by maximum likelihood estimators, then derive the optimal pricing policy and its properties. Since the exact optimal policy is difficult to implement for problems in practical scale, we propose two simple and computationally tractable pricing policies that are provably near-optimal. Our framework is sufficiently flexible to be applied to real-time control optimization and parameter learning with general network effect in demand process.

In the second part of the dissertation, we consider a setting when underlying demand is uncertain and follows a complex stochastic process, which makes pricing problems difficult to solve. In such cases, certainty equivalent (CE) policies, based on solving the deterministic relaxation of a stochastic pricing problem, can be used as practical alternatives. CE policies have lighter computational and informational requirements compared to solving the problem to optimality. This is particularly true when the firm does not have complete information about the underlying demand distribution.

While the effectiveness of CE pricing policies has been theoretically studied in some settings (e.g, independent demand), the performance of CE policies are not known in general settings. This paper analyzes the performance of CE policies in a pricing problem (for a given inventory level) where future demand depends on sales and inventory and the firm has limited opportunities to change price. We show that CE policies are asymptotically optimal: as the problem scale (denoted by m) becomes large, the percentage regret decreases at the rate of $\Theta(1/\sqrt{m})$. We also extend the result to the joint pricing and (initial) inventory problem. Our numerical results are even more promising. Even in non-asymptotic settings (small scaling factor and a few price changes), CE policies perform well and often result in revenues that are only a few percentage points lower than optimal.

The third part of the dissertation is an on-going project which studies the cooperated relationship between a seller and an social influencer where the influencer has a revenue-sharing contract with the seller (so that their incentives are aligned). We study the following questions: (*i*) for the seller, how to compensate the influencer? (*ii*) Who is the ideal influencer to post about a given product? (*iii*) What attributes of an influencer make her more or less valuable to a company? The customers are assumed to be Bayesian learners so they can infer the state of the product from available information.

CHAPTER I

Introduction

Traditional models in revenue management commonly have three simplifying assumptions: (*i*) the firm has a large certainty about the product demand model; (*ii*) customers are isolated entities that independently make their decisions; and (*iii*) the firm has the ability to continuously change the price so that each customers can receive different prices. In this dissertation, we focus on revenue management practices in the settings where the previous three assumptions are challenged because of the rise of social technologies in the new age. We develop theoretically sound policies (pricing, inventory, and information signals) when the firm faces uncertainty, dependency and non-stationarity in market demand. We also provides effective insights into the firm's decision making.

The second chapter is devoted to the demand learning and pricing problem for a new product. Decisions regarding new products are often difficult to make, and mistakes can have grave consequences to a firm's bottom line. However, firms often have little foresight on important information about new product demand such as the potential market size, the rate of customers' adoption, and their willingness to pay. One of the most popular frameworks that have been used for modeling new product adoption is the Bass model (*Bass*, 1969). While the original Bass model and its many variants are useful to understand the factors and decisions affecting new product adoptions over time, the vast majority of these models require *a priori* knowledge of key parameters, which can only be estimated with historical data or guessed based on institutional knowledge.

As is found by *Bass* (1969), there exist significant dependency and non-stationary (e.g., word-of-mouth effects) in new product adoption processes. Therefore, the current price not only affects the current revenue, but also the number of adopters who can influence sales for the product in future periods. In addition to the dynamics of current price and future demand, pricing decisions (initial price and subsequential price changes) for a new product are challenging due to the limited sales data available. Due to the lack of historic data, it is difficult to estimate how the future demand will respond to a price change. Therefore, any tool that enables a firm to enhance its capability in forecasting and pricing of new products will be highly valuable.

In this chapter, we study the interplay between pricing and learning for a monopolist whose objective is to maximize the expected revenue of a new product over a finite time horizon. We consider a setting where a firm can learn by observing sales data at (different) prices over time. To capture the adoption process with demand learning, we develop a continuous-time Markovian Bass model where adoption rate depends on a selling price and on the past sales. We first show that the loss from a wrong initial demand estimation can grow proportionally in time and can be significant. We then show that a firm can apply real-time demand data to update the parameter and optimize the price accordingly. We formulate the problem as a stochastic optimal control problem where the demand parameters are updated by maximum likelihood estimators, then derive the optimal pricing policy and its properties. Since the exact optimal policy is difficult to implement for problems in practical scale, we propose two simple and computationally tractable pricing policies (continuous price change and limited number of price changes) that are provably near-optimal. Our framework is also sufficiently flexible to be applied to real-time control optimization and parameter learning with general network effect in demand process.

The third chapter is motivated by the difficulty of designing (dynamic) pricing policies when demand is uncertain and depends on factors that can change over time. For example, when future demand is driven by a network effect, 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. In order to determine the optimal prices in such settings, the seller must know the distribution of future demand. However, when demand is a complex and state-varying stochastic process, the seller may not have complete demand information. In many settings, the seller's best available information is the estimate of the average demand in future periods. Indeed, estimating conditional means from data uses standard statistical methodologies (relying on strong results like the law of large numbers), whereas estimating an entire distribution requires a much larger data set and more sophisticated approaches.

In this chapter, we study a periodic-review[1] pricing problem over a finite horizon and

[1]Periodic review means that prices can only be changed at the start of each period. While a continuous review of pricing is ubiquitous in analytical models of dynamic pricing, periodic pricing changes are often more appropriate in reality (*Yang and Zhang*, 2014; *Bitran and Mondschein*, 1997). Indeed, periodic

with finite inventory when the demand distribution is state-varying. The key features that distinguish our demand model from others in the dynamic pricing literature are that we assume that the future demand and its distribution are state-dependent (where the state variables in our setting are the total past sales and the current inventory level) and that the seller only has limited information about the demand distribution. When demand is state-dependent, a pricing mistake not only reduces the current period revenue, but also changes future demand since the mistake affects cumulative sales and available inventory. Thus, price in one period has a lingering effect on future demand. Furthermore, when there is a limited number of opportunities to change price, the price chosen at each period has persistent implications beyond the current period. Lack of knowledge about the demand distribution makes the pricing decision more difficult and nonoptimal pricing more consequential.

We study a class of pricing policies called certainty equivalent (CE) pricing policies. Certainty equivalent pricing policies are commonly used when the seller has access to the expected demand rather than the entire distribution. Specifically, 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 (CE-OL) implements the optimal price sequence of the deterministic model. Although 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 (CE-CL) re-optimizes the deterministic model on a rolling horizon using current inventory information at the beginning of each period. Hence, prices are adjusted over time based on the realizations of demands in past periods.

We show that as the scaling factor m increases, both CE pricing policies are asymptotically optimal with a regret rate of $\Theta(\sqrt{m})$, when compared with the optimal policy. However, in a non-asymptotic setting, CE-CL performs consistently better than CE-OL, especially when the number of price change periods increases and when the conditional expectation of demand is highly nonlinear in past cumulative sales and available inventory. This result highlights the importance of re-optimization in the face of sales and inventory dependent demand. We then extend our results to the case where the seller chooses initial inventory along with 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.

In the last chapter, we depart from developing operations decisions (pricing or inven-

pricing schemes are widely observed in practice. For example, many brick-and-mortar stores update their prices weekly as changing prices often requires changing price stickers and are costly to implement.

tory control) that *utilize* demand hype, by considering the best informational signaling strategies that *creates* demand hype. We propose a model where the company (a seller) hires a social influencer at a cost. *Influencer marketing* — promoting products or services through social-media *influencers* — has been a popular practice in recent years. Influencers are individuals who share their impressions of a particular product category (such as fashion, technology, gaming, travel) in order to shape the opinions of their followers. They create content on social media platforms (e.g., YouTube, Instagram, and Facebook) with active followers who consume this content. Influencers earn money through their reputation, which includes being consistent and true to their values. Companies leverage these well-earned reputations in exchange for money, to target customers who follow influencers. Companies recognize the value of influencers in spreading valuable information about their products, particularly information that they may not be able to credibly communicate themselves. This marketing practice allows companies to attract interest in a product from pre-selected active customers. According to a report from Wondershare (*Brown*, 2020), a platform providing video-making software, the top ten richest content creators on YouTube earned between \$10 million to \$15.5 million in 2019. InfluencerDB, a platform that collects data on influencer marketing, reports that spending on Instagram influencers alone exceeded \$5 billion in 2018 (*Vardhman*, 2019).

Although companies spend large sums on influencer marketing, 61% of them find it difficult to identify effective influencers for their product campaign *MediaKix* (2019). An effective influencer increases brand awareness and converts awareness into sales. Influencers are typically paid by companies by a flat rate per “post”, where a post is a video, picture, or message on a social media platform. Revenue-sharing arrangements exist but are much rarer in practice.[2] However, the *value* of a post is hard to assess. The following questions are still largely unanswered:

(i) how lucrative is it for a company to work with a given influencer?

(ii) who is the ideal influencer to post about a given product?

(iii) what attributes of an influencer make her more or less valuable to a company?

This chapter develops an analytical model designed to answer these and related questions.

[2]Verified through personal communication with social media influencers.

CHAPTER II

Data-driven Pricing for a New Product

II.1 Introduction

Decisions regarding new products are difficult and risky because mistakes can have grave consequences for a firm's bottom line. Before a product launch (or even after a launch), firms often have little information regarding demand (such as the market size or the speed of adoption by customers). The lack of information makes pricing a new product very challenging.

To cope with this information deficiency, firms use several strategies to forecast the demand for a new product. For instance, historical data of *similar* products could be used to infer the new product's demand characteristics. Alternative techniques include forecasting based on judgment (e.g., using an expert opinion) or market research (*Kahn*, 2006). Yet, due to insufficient or inaccurate data, the expense of market research or subjective biases injected by management, forecasting new product demand is prone to errors. Such errors lead to weak market penetration (due to a market price that is too high) or to lost potential revenue (due to a market price that is too low).

For instance, Apple released its new generation of smartphones, iPhone XS (priced at \$999), iPhone XS Max (\$1,100) and iPhone XR (\$749, which is the "budget" choice to replace the previous \$349 SE model) in 2018. Even though the new iPhones received many technological improvements, many consumers found it difficult to accept a \$1,100 price tag (*USA TODAY*, 2019). In Q1 of 2019, iPhone's revenue declined 15 percent from the previous year's even though, in that same quarter, revenue for Apple's other products and services grew significantly (*Business Wire*, 2019). In an earnings call, Apple's CEO Tim Cook admitted that, though the weakened U.S. dollar in Q1 2019 accounted for some of this decline, "price was a factor" for the iPhone's weak performance since the cheapest model (iPhone XR) was also the most popular among the new models (*ZDNet*, 2019).

Another example can be found in high-end streetwear brands such as Yeezy and Supreme. These brands release new styles or colors of sneakers sporadically. The sneakers are usually sold for a limited time and in limited quantities. As soon as the new products are released, celebrities, influencers and collectors start to create a buzz for some styles on social media platforms. The resale price of these new styles can have a markup as high as 1000% compared with the original release price (*BBC*, 2018). Judging by the aftermarket sales, streetwear brands can severely underestimate consumers' valuation, inadvertently leaving a significant portion of the revenue on the table.

These examples illustrate how hard it is to price a new product. First of all, pricing decisions (about the initial price and subsequent price changes) for a new product are challenging due to the limited sales data available. As a result, it is difficult to estimate how the market will respond to a price change. Furthermore, the demand for a new product is often influenced by how many people have bought the product so far, which creates the word-of-mouth effect. Thus, the current selling price not only affects the current revenue and demand, but it also influences how quickly product adoption will occur in the future.

In this chapter, we study the interplay between demand learning and dynamic pricing for a new product. We propose an analytical framework that captures parameter learning while pricing during the new product adoption for a monopolistic firm. In order to model the dynamics of demand and learning in a tractable way that is consistent with existing literature on new product adoption, we modify the generalized Bass model (*Bass et al.*, 1994) to capture stochastic adoption in a *Markovian Bass model*. The main contribution of this paper is outlined below.

- *Markovian Bass model.* Traditional stochastic adoption models that add noise to the cumulative demand are not well-behaved when modeling new product adoption. For instance, Brownian diffusion models violate the fact that cumulative sales must be nondecreasing in time. To overcome this technical challenge, we propose a different way to introduce stochasticity in an adoption process while capturing the features of the Bass model. We model the cumulative adoption process as a continuous-time Markov chain where the time between adoptions depends on price and cumulative sales. We refer to this as the *Markovian Bass model*. We show that the Markovian Bass model converges to the original Bass model as the market size grows. Further, we derive the optimal pricing policy (MBP) under a Markovian Bass model when the seller has complete information—this setting is used as a benchmark when evaluating data-driven pricing policies.

- *Demand learning.* We establish several theoretical properties of the maximum likelihood (ML) estimators of the Markovian Bass model parameters. First, we derive sufficient conditions for the parameters to be identifiable. Second, using Itô calculus, we derive the stochastic differential equations that govern the dynamics of the ML estimators of this demand model. Third, we establish the convergence properties of the estimation error of the ML estimators. The challenge in proving the third result is that the data on inter-adoption times is non-*i.i.d.*, so we cannot use the standard proof techniques to show convergence when data is *i.i.d.*. We circumvent this impediment and show that, in this non-*i.i.d.* setting, the mean squared errors of the ML estimators are inversely proportional to the number of adopters.
- *Optimal learning and pricing policy.* We formulate the seller's pricing-and-learning problem when the seller forms an inference about demand using MLE. Interestingly, we show is that the MLE dynamics in this control problem has a Markov property. This enables a state reduction for the control problem. We derive the Hamilton-Jacobi-Bellman (HJB) equation of the control problem and characterize the optimal pricing-and-learning policy.
- *Performance guarantee for tractable pricing policies.* Due to the computational challenges of computing the optimal pricing-and-learning policy, we propose two computationally efficient data-driven pricing policies. The first policy (MBP-MLE) is a tractable approximation of the optimal MBP policy and can be used in a setting where a firm can change the price frequently. The second policy (MBP-MLE-Limited) reflects a business constraint that the firm can change prices a limited number of times. We provide analytic performance bounds for both policies and show each has a worst-case regret that is in the order of the log of the market size. We prove a fundamental lower bound to the regret of any pricing-and-learning policy, and show that the regret of our policies matches this limit.

II.1.1 Review of related literature

This chapter is related to the literature on new product adoption models as well as dynamic pricing and learning. Both areas draw on a considerable body of literature from economics, marketing, and operations research. We also review the literature on continuous-time Markov chains (CTMCs) with unknown transition rates since our Markovian Bass model is essentially a CTMC.

II.1.1.1 New product adoption models.

Since the seminal work by *Bass* (1969), numerous papers have used a Bass-like model to explain new product adoption. In the original Bass model, sales are temporally influenced by innovators (who try a product on their own) and imitators (who follow earlier adopters). Variants of the Bass model have been used to explain the impact of competition (*Krishnan et al.*, 2000; *Savin and Terwiesch*, 2005; *Guseo and Mortarino*, 2013) and of overlapping generations (*Norton and Bass*, 1987; *Bayus*, 1992). Comprehensive surveys of adoption models are provided by *Mahajan et al.* (1995) and *Baptista* (1999). Most relevant to our work are the variants that explain the role of price in adoption, such as the generalized Bass model introduced in *Bass et al.* (1994). There has been a long tradition in marketing literature to derive the optimal pricing policies for new products under variants of the Bass model (*Robinson and Lakhani*, 1975; *Dolan and Jeuland*, 1981; *Bass and Bultez*, 1982; *Kalish*, 1983; *Horsky*, 1990; *Krishnan et al.*, 1999). Dynamic pricing under a Bass-type model has also recently gained attention in operations (*Li and Huh*, 2012; *Shen et al.*, 2013; *Li*, 2020). However, most of these works assume deterministic adoptions.

Raman and Chatterjee (1995) and *Kamrad et al.* (2005) study pricing under a stochastic adoption process by adding a normally distributed noise to the Bass adoption rate. While adding Brownian noise can leverage stochastic calculus, Brownian noise violates the fact that cumulative sales must be non-decreasing in time. Alternatively, *Böker* (1987) and *Niu* (2002) propose modeling adoption as a counting process, though these works do not consider the fact that a firm can influence adoption through pricing decisions. Our model uses a counting process as a model construct in a setting where a firm can dynamically control price. Furthermore, none of the aforementioned works (deterministic or stochastic) study the interplay between pricing and learning.

In both stochastic and deterministic models, a common assumption is that the firm knows the key parameters of the demand model. These parameters include the market size (denoted by m_0), the innovation rate (p_0), and the imitation rate (q_0). In the case of unknown parameters, *Bass* (1969) and *Srinivasan and Mason* (1986) propose least squares methods to estimate these parameters, whereas *Schmittlein and Mahajan* (1982) suggests using maximum likelihood estimation. However, these approaches assume that the firm has sufficient data to build accurate estimates. Our model uses the stochastic Bass model for pricing decisions during the product launch, hence we study how a revenue-maximizing firm, which starts with very little demand information, can use pricing and real-time data to better calibrate the demand parameters.

II.1.1.2 Dynamic pricing and learning.

There is a growing literature on dynamic pricing with limited demand information (see surveys in *Araman and Caldentey* 2010 and *den Boer* 2015a). Some papers use parametric approaches in demand learning. These papers assume that an underlying model belongs to a parametric family and the unknown parameters are estimated using various estimators. *Lin* (2006); *Araman and Caldentey* (2009); *Farias and Van Roy* (2010) and *Harrison et al.* (2012) study Bayesian learning. Other learning methods include regression (*Bertsimas and Perakis*, 2006) and maximum likelihood estimation (*Besbes and Zeevi*, 2009; *Broder and Rusmevichientong*, 2012; *Keskin and Zeevi*, 2014; *den Boer and Zwart*, 2015). On the other hand, nonparametric approaches do not impose a particular form to model underlying demand. *Lim and Shanthikumar* (2007); *Besbes and Zeevi* (2009) and *Eren and Maglaras* (2010) use the worst-case analysis to develop robust policies. *Kleywegt et al.* (2002) use sample average approximation to approximate underlying demand. *Ferreira et al.* (2015) consider a price optimization model where the demand information is estimated with a regression tree.

Under a Markovian Bass model, the market is nonstationary because the adoption rate depends on how many customers have already adopted. Dynamic pricing with demand learning in a time-varying market is largely unexplored. *Besbes and Zeevi* (2011) and *Besbes and Sauré* (2014) consider settings where the willingness-to-pay distribution changes at some unknown time. *Keskin and Zeevi* (2016) study an unknown time-varying demand with a constraint on the number of price changes. *Chen and Farias* (2013) and *den Boer* (2015b) study pricing policies under a setting where the time-varying market size is unknown. Our work adds to this literature by studying the pricing strategies under an adoption model where the unknown time-varying demand rate is influenced by the price and by the changing cumulative adoptions.

II.1.1.3 Learning in stochastic processes.

Our work is related to estimating the unknown transition rates in continuous-time Markov chains (CTMCs). *Duffie and Glynn* (2004) propose a family of generalized-method-of-moments (GMM) estimators sampled at random time intervals. On the other hand, *Kessler* (1995, 1997, 2000) considered GMM estimators using data samples taken at deterministic time intervals (discrete observations). Although all those estimators are consistent, GMM methods are more computationally challenging than MLE methods, which are derived from the first-order conditions. There are other approaches, which include simulation-based methods (e.g., the simulated-method-of-moments estimation studied by

Duffie and Singleton 1993) and nonparametric estimations (e.g., approximating transition rates using analytic expansions studied by *Aït-Sahalia* 2002). However, theoretical results with these approaches are only limited to cases where the random noise follows a Brownian motion (or one of its variants).

There exists literature addressing optimal controls under the setting where the transition matrix of a Markov decision process is unknown. For example, *Araman and Caldentey* (2010) propose a Bayesian approach to learn an unknown parameter of a price-modulated Poisson process. A Bayesian method requires the knowledge of the prior distribution; further, it is time-consuming to compute when there are multiple unknown parameters to learn. Several papers (*Nilim and El Ghaoui*, 2005; *Kalyanasundaram et al.*, 2002; *Nilim and El Ghaoui*, 2004) consider robust control problems for Markov decision processes with unknown and stationary transition matrices. Our estimation, on the other hand, is based on MLE and uses the first-order conditions. We can analytically characterize the updating processes of the MLEs used in the Markovian Bass model. Our method is amenable to the case where there are multiple unknown parameters. In fact, the main results of our paper hold when the firm does not know the market size, the adoption innovation rate, the imitation rate, and the price sensitivity function. We contribute to the literature on learning and controls of CTMCs by proposing a maximum likelihood approach to a Markov decision process where transition rates evolve as more adoptions occur.

II.1.2 Preliminaries

In the paper, we use the big $\mathcal{O}$ notation where, by definition, $f(x) = \mathcal{O}(g(x))$ for positive real-valued functions f and g if there exists an $r \in \mathbb{R}$ such that $f(x) < rg(x)$. 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 $f(x) = \Theta(g(x))$.

II.2 The model

We first discuss the stochastic demand model of new production adoption. Then, we formally state the seller's pricing-and-learning problem.

II.2.1 The stochastic demand model

Bass (1969) proposed a model for the timing of adoptions of a new product, where the adoption rate increases with the number of past adoptions. There have since been numerous extensions of this model. One notable extension relevant to our work is the generalized Bass model (*Bass et al.*, 1994; *Krishnan et al.*, 1999) where price influences adoptions. We first review the generalized Bass model and establish model constructs for the adoption model we will use in the paper.

The generalized Bass model represents adoption timings of a new product by a market of customers under a known price path. Let $r = \{r_t, t \geq 0\}$ denote the price sequence where r_t represents the price at time t, where $r_t \in (-\infty, \infty)$.[1] Given this price path, let F_t^r be the proportion of the market that has adopted the product by time t, where $F_t^r \in [0, 1]$. In the case where F_t^r is continuously differentiable in t, then $f_t^r = \frac{\mathrm{d}F_t^r}{\mathrm{d}t}$ is the marginal rate of adoption and $f_t^r/(1 - F_t^r)$ is its failure rate. The generalized Bass model assumes that the time t failure rate (i.e., the marginal rate of adoptions among the remaining customers at time t) is equal to $(p_0 + q_0 F_t^r)x(r_t)$, where $x(\cdot)$ is a *marketing effort* function that reflects the effect of price. Here, p_0 (where $p_0 > 0$) is called the coefficient of innovation and it represents the rate at which consumers adopt the product on their own initiative. On the other hand, q_0 (where $q_0 > 0$) represents the imitation coefficient, representing the rate at which consumers imitate earlier adopters (through word-of-mouth effect or a network effect). Note that the firm can influence the adoption process by setting the price sequence r.

The cumulative adoption proportion F_t^r satisfies the following differential equation:

$$\frac{\mathrm{d}F_t^r}{\mathrm{d}t} = (1 - F_t^r)(p_0 + q_0 F_t^r)x(r_t). \tag{II.2.1}$$

We can compute its solution as

$$F_t^r := \frac{1 - e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}}{1 + \frac{q_0}{p_0} e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}}. \tag{II.2.2}$$

The generalized Bass model (II.2.1) assumes that the adoptions are deterministic with an adoption function F_t^r. While deterministic models are useful in understanding the trajectory of expected adoption over time under a given price path, they fail to model random choices of individual customers and their impact on overall adoptions. A few papers propose stochastic adoption models (*Raman and Chatterjee*, 1995; *Kamrad et al.*, 2005), yet these papers use Brownian models that fail to enforce that the cumulative adoption is a non-decreasing process. We follow a different approach by assuming that the time between two successive adoptions is random. The result is a counting process

[1] In the new product pricing literature, price is allowed to be zero or negative. This is because it might be beneficial for the seller to offer the product for free or even compensate early adopters in order to increase the future adoption. See *Kalish* (1983) and *Krishnan et al.* (1999) for example.

which we refer to as the *Markovian Bass model* because, inspired by the Bass model, it obeys the Markov property.

We define $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}\}_{t\geq 0})$ as a filtered probability space endowed with a cumulative adoption process $D = \{D_t,\ t \geq 0\}$ where D_t is the cumulative adoptions by time t. Let m_0 be a positive integer that denotes the market size of potential customers. Hence, $D_t : \Omega \mapsto \{0, 1, \ldots, m_0\}$. Since adoptions can only occur in unit increments, D is a counting process. Let $\{\mathcal{F}_t,\ t \geq 0\}$ be the history or filtration associated with the process of prices and adoptions, with $\mathcal{F}_t = \sigma((r_s, D_s), s \in [0, t])$. We say that π is a non-anticipating pricing policy if the price r_t^π offered by π at time t is $\mathcal{F}_t$-measurable. If customers are price-sensitive, a price change results in a change in the adoption rate. To explicitly state the dependence in price, we will henceforth refer to the cumulative adoption as D^π instead of D. Without loss of generality, we assume that $D_0^\pi = 0$ for any π, thus none of the consumers purchases before time $t = 0$.

As in the Bass model, the adoption rate in the Markovian Bass model is also dependent on a coefficient of innovation, p_0, and a coefficient of imitation, q_0. We denote the parameters of the Markovian Bass model as $\theta_0 := (p_0, q_0, m_0)$, where $p_0, q_0 > 0$. If at time t, the cumulative number of adoptions is j and the seller sets price r_t, then under the Markovian Bass model the transition rate to the next $(j+1)$-st adoption is

$$\lambda(j, r_t) := \xi(j) \cdot x(r_t), \quad \text{for } j = 0, 1, \ldots, m_0, \tag{II.2.3}$$

where

$$\xi(j) := (m_0 - j)\left(p_0 + q_0 \cdot \frac{j}{m_0}\right). \tag{II.2.4}$$

Note that $\xi(j)$ is the portion of the adoption rate unaffected by price. From (II.2.4), we see that each of the $m_0 - j$ potential adopters are homogeneously affected by their own will to adopt the product (reflected in term p_0) and by the influence from previous adopters (reflected in term $q_0 \frac{j}{m_0}$). We will sometimes write $\xi(j; \theta_0)$ or $\lambda(j, r_t; \theta_0)$ to emphasize the dependence of these values on θ_0. Given the pricing policy π, the adoption process is a nonhomogeneous, continuous-time Markov chain with the following transition probabilities. For a small time interval of size h,

$$\mathbb{P}_{\theta_0}\left(D^\pi_{t+h} = j + k \mid D^\pi_t = j\right) = \begin{cases} 1 - \lambda(j, r^\pi_t)h + o(h), & \text{if } k = 0, \\ \lambda(j, r^\pi_t)h + o(h), & \text{if } k = 1, \\ o(h), & \text{if } k \geq 2, \end{cases} \tag{II.2.5}$$

where $o(h)$ is a term such that $\lim_{h\to 0} o(h)/h = 0$. The subscript θ_0 on $\mathbb{P}_{\theta_0}$ is to denote the dependence of the probability on the parameter vector θ_0. Note that the Markovian Bass model guarantees that the cumulative adoption is always non-decreasing.

Conditional on $\mathcal{F}_t$, the expected demand rate is

$$\mathbb{E}_{\theta_0}\left[\mathrm{d}D^\pi_t \mid \mathcal{F}_t\right] = \lambda(D^\pi_t, r^\pi_t)\mathrm{d}t = (m_0 - D^\pi_t)\left(p_0 + q_0 \cdot \frac{D^\pi_t}{m_0}\right) x(r^\pi_t)\mathrm{d}t. \tag{II.2.6}$$

Hence, the Markovian Bass model captures the demand dynamics in the generalized Bass model (II.2.1). First, the expected demand rate is increasing in the remaining market size, $m_0 - j$, and decreasing in price, r_t. Second, adoptions occur naturally or imitatively. As in the generalized Bass model, the rate of adoption also depends on the proportion of customers who have adopted, D^π_t/m_0. By including a price effect, the Markovian Bass model generalizes the stochastic Bass model proposed by *Niu* (2002, 2006). Since the price process r^π is endogenous, this seemingly innocuous extension has implications in the convergence results and their proofs.

We state a property of the evolution of Markovian Bass model that is consistent with the generalized Bass model studied in *Bass et al.* (1994) and *Krishnan et al.* (1999).

Lemma II.1. The increase in the adoption speed decreases as more people have adopted the product. In other words, for a given price r, $\lambda(d, r) = (m_0 - d)(p_0 + q_0 \frac{d}{m_0})x(r)$ is concave in d.

Next, we show that the Markovian Bass model is consistent with its deterministic counterpart under a deterministic price process r. Let $\{D^{r,m_0},\ m_0 \geq 1\}$ be the family of Markovian Bass models indexed by market size m_0. As m_0 increases, we show that the proportion of customers who have purchased by time t (the cumulative proportion) converges to the deterministic Bass curve.

Proposition II.1. For a given price sample path $r = \{r_t,\ t \geq 0\}$, if $\{D^{r,m_0}_t,\ t \geq 0\}$ is the cumulative adoption process with market potential m_0, then the following holds

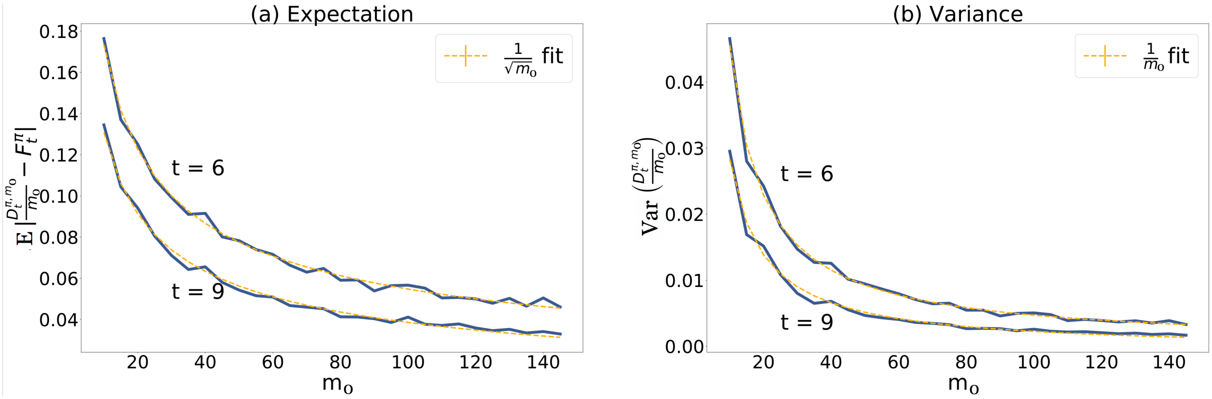


Figure II.1: Convergence of expectation and variance of $D_t^{\pi,m_0}/m_0$ as m_0 increases.

uniformly in t:

$$\frac{D_t^{r,m_0}}{m_0} \to F_t^r \text{ almost surely as } m_0 \to \infty, \tag{II.2.7}$$

where F_t^r is given by (II.2.2), and $\mathrm{Var}_{\theta_0}\left(\frac{D_t^{r,m_0}}{m_0}\right)$ decreases in the order of $\mathcal{O}({m_0}^{-1})$.

Proposition II.1 states that the variance of $\frac{D_t^{r,m_0}}{m_0}$ diminishes to zero at the rate that is inversely proportional to m_0. Additionally, we have Lemma II.2 below, which shows that the expectation converges to F_t^r at a rate inversely proportional to $\sqrt{m_0}$. These results are helpful in understanding the behavior of a Markovian Bass model at an asymptotic regime.

Lemma II.2. For a given price sample path $r = \{r_t,\ t \geq 0\}$, if $\{D_t^{r,m_0},\ t \geq 0\}$ is the cumulative adoption process with market size m_0, then

$$\mathbb{E}_{\theta_0}\left|\frac{D_t^{r,m_0}}{m_0} - F_t^r\right| = \mathcal{O}\left(\frac{1}{\sqrt{m_0}}\right) \quad \text{for all } t > 0. \tag{II.2.8}$$

Figure II.1 illustrates Lemma II.2 and Proposition II.1 by showing how the proportion of adopters by time t, $\frac{D_t^{r,m_0}}{m_0}$ behaves. Panel (a) shows how the difference of the expected proportion from F_t^r changes with increasing m_0. Panel (b) shows how the variance of the adoption fraction changes with increasing m_0. We compute the expectation and variance by simulating 10^3 sample paths of the adoption process with $p_0 = 0.1,\ q_0 = 0.3,\ r_t = 0.1 + \frac{t}{100}$, and $x(r) = e^{-r}$. We observe that the expected difference between the Markovian Bass adoption, $\frac{D_t^{r,m_0}}{m_0}$, and the deterministic Bass adoption, F_t^r, decreases in the order of $\frac{1}{\sqrt{m_0}}$ and the variance decreases in the order of $\frac{1}{m_0}$.

II.2.2 Seller's pricing-and-learning problem

We consider the dynamic pricing and learning problem of a monopolist launching a new product over a finite selling horizon $[0, T]$ where $T > 0$. In this setting, the demand for the new product is described by the Markovian Bass model with parameters $\theta_0 = (p_0, q_0, m_0)$, but the seller does not know the parameters of the underlying demand model. However, the seller can accumulate market information, represented in $\mathcal{F}_t$, by continuously monitoring its prices and sales throughout the selling horizon. The seller uses statistical inference on the observed demand data in order to infer the unknown demand parameters.

Our goal is to understand how a firm can use the price and sales data after a product launch to learn the true characteristics of the underlying demand model. In particular, the seller's problem is to dynamically adjust the prices to maximize the expected cumulative revenue by utilizing what the seller learns about the demand parameters from data collected over time.

We denote the set of non-anticipating pricing policies as Π, which is the set of $\mathcal{F}_t$-adapted pricing policies. When θ_0 is unknown, the optimal pricing-and-learning policy is the solution to:

$$\sup_{\pi\in\Pi} R(\pi) := \sup_{\pi\in\Pi} \mathbb{E}_{D|\mathcal{F}_0}\left[\int_0^T r_t^\pi \, \mathrm{d}D_t^\pi \mid \mathcal{F}_0\right] = \sup_{\pi\in\Pi} \mathbb{E}_{D|\mathcal{F}_0}\left[\int_0^T \mathbb{E}_{D|\mathcal{F}_t}\left[r_t^\pi \, \mathrm{d}D_t^\pi \mid \mathcal{F}_t\right] \mid \mathcal{F}_0\right]. \tag{II.2.9}$$

The goal is to find a policy π that maximizes the expected total revenue $R(\pi)$. Since (II.2.9) is solved by the seller, the expectation cannot rely on the true demand parameters θ_0 that are unknown to the seller. Hence, $\mathbb{E}_{D|\mathcal{F}_t}$ in (II.2.9) is the conditional expectation given the information set $\mathcal{F}_t$ known to the seller at time t. The demand follows a Markovian Bass model with unknown parameters, so the expectation is with respect to the seller's inference on those parameters given $\mathcal{F}_t$.

There are two schools of statistical inference: Bayesian and frequentist. If the seller forms an inference using Bayesian learning, then conditional on $\mathcal{F}_t$, D_t^π follows a Markovian Bass model with a *random* parameter vector θ. The posterior distribution of θ is computed using Bayes' rule by updating the prior distribution given the observed data. On the other hand, if the seller forms an inference using a frequentist approach, then conditional on $\mathcal{F}_t$, D_t^π follows a Markovian Bass model with parameter vector $\hat{\theta}_t$, where $\hat{\theta}_t$ is a point estimate (e.g. maximum likelihood estimator) computed from the data. The frequentist school is rooted in the philosophy that there is only one true value for the parameters, so it does not make sense to assign probabilities (prior or posterior) for the different parameter values. Section A.3 in the Online Appendix discusses how to formulate the value functions of (II.2.9) under Bayesian inference and frequentist inference, respectively.

Price r_t^π plays two roles in this setting. The first role is to affect the revenue and demand at time t. Note that in the Markovian Bass model, the demand at time t has a compounding effect since it influences the probability of future purchases through the imitation effect. The second role is to affect the price and sales information that will be used for demand inference in future periods. Due to these two roles of price, we refer to (II.2.9) as the seller's pricing-and-learning problem.

The Bass diffusion model has a long tradition in marketing literature of being used in deriving optimal dynamic pricing policies, starting from *Robinson and Lakhani* (1975). Some early examples of dynamic pricing under the Bass model include *Dolan and Jeuland* (1981); *Bass and Bultez* (1982); *Kalish* (1983); *Horsky* (1990); *Raman and Chatterjee* (1995); *Krishnan et al.* (1999). More recent examples in the operations literature are *Kamrad et al.* (2005); *Li and Huh* (2012); *Shen et al.* (2013); *Li* (2020). In this paper, we continue this tradition by studying how learning affects the pricing decisions under a stochastic version of the generalized Bass model.

To guarantee the existence of a unique optimal price for the control problem (II.2.9), we assume that the marketing effort function $x(\cdot)$ satisfies certain regularity properties.

Assumption II.1. The marketing effort function $x : \mathbb{R} \to \mathbb{R}^+$ has the following properties:

i. (Smoothness and bounded derivative.) x is twice differentiable, and there exists $M > 0$ such that $|x'(r)| \leq M$, for all $-\infty < r < \infty$;
ii. (Non-negativity.) There exist non-negative constants $\bar{x}^u, \bar{x}^l$, such that $\bar{x}^l t \leq \int_0^t x(r_s)\mathrm{d}s \leq \bar{x}^u t$, for all $r = (r_s, s \geq 0)$ where $-\infty < r_s < \infty$ for all $s \geq 0$;
iii. (Decreasing in price.) $x'(r) < 0$ for all $-\infty < r < \infty$;
iv. (Monotone hazard rate.) $r + \frac{x(r)}{x'(r)} + C$ is strictly monotone increasing in r with a finite root for any finite C, and for all $-\infty < r < \infty$, there exists a constant $C_d > 0$, such that $2x'(r)^2 - x(r)x''(r) \geq C_d > 0$;
v. (Boundedness of revenue.) There exist constants C_x, C_{xx} such that, for $f(r) := rx(r)$, $|f(r)| \leq C_x$, and $|f''(r)| \leq C_{xx}$ for all $-\infty < r < \infty$.

Assumption II.1(i)–(iii) are innocuous as they guarantee that the market effort function is decreasing in price and is sufficiently smooth. Assumption II.1(iv) is a standard assumption to ensure that the revenue function is well-behaved and has a unique optimal price for a given state. Assumption II.1(v) implies the revenue function is bounded. The bounded second-order derivative is an assumption used in many papers (*Broder and Rusmevichientong*, 2012; *Wang et al.*, 2014). These properties are satisfied by many functional forms including multiplicative (e.g., $x(r) = e^{a-br}$), and additive (e.g., $x(r) = a - br$) relationships.

In the following sections, we study (II.2.9). First, we study the problem assuming that the seller knows the parameter vector θ_0 that governs the adoption process. Then, we study the pricing-and-learning problem where the seller forms her inference about the unknown parameters from past data through maximum likelihood estimation.

II.3 Optimal pricing policy with complete information

If the firm knows the demand model parameter vector θ_0, the optimal control problem is

$$R^* := \sup_{\pi \in \Pi} \ \mathbb{E}_{\theta_0}\left[\int_0^T r_t^\pi \ \mathrm{d}D_t^\pi\right] = \sup_{\pi \in \Pi} \ \mathbb{E}_{\theta_0}\left[\int_0^T \mathbb{E}_{\theta_0}\left[r_t^\pi \ \mathrm{d}D_t^\pi \ \mid \ \mathcal{F}_t\right] \ \middle| \ \mathcal{F}_0\right], \qquad \text{(II.3.1)}$$

where the equality follows from the tower property of conditional expectation. Note the distinction of (II.3.1) from the pricing-and-learning problem (II.2.9). In (II.3.1), the seller knows θ_0, so the expectation $\mathbb{E}_{\theta_0}$ is taken with respect to the Markovian Bass model with parameter vector θ_0.

We denote the optimal expected cumulative revenue under complete information as R^*. Note that R^* is the expected revenue of an oracle that knows the true value of θ_0, so it is an upper bound to the optimal expected revenue in the pricing-and-learning problem (II.2.9).

To solve the optimal control problem (II.3.1), we define $V(d, t)$ to be the optimal value-to-go function where t is the time remaining until the end of the horizon T, and d is the cumulative number of adoptions after $T - t$ time has elapsed. Hence,

$$\begin{aligned} V(d,t) \ := \underset{\pi \in \Pi}{\text{maximize}} \quad & \mathbb{E}_{\theta_0}\left[\int_{T-t}^T r_s^\pi \ \mathrm{d}D_s^\pi\right] \\ \text{subject to} \quad & D_{T-t}^\pi = d. \end{aligned}$$

Note that $V(0, T)$ is the optimal expected revenue of the pricing problem (II.3.1).

We will sometimes write $V(d, T; \theta_0)$ to emphasize that the value function depends on the demand parameters θ_0. This will prove useful in later sections when θ_0 could be replaced by a data-driven estimator. From (II.2.4), we know $\theta_0 = (p_0, q_0, m_0)$ affects the optimal expected revenue through its effect on the adoption rate. Therefore, increasing p_0, q_0 or m_0 results in a higher adoption rate, and consequently, a higher expected revenue. This is formally stated in Lemma A.1(iii).

We can write $V(d,t)$ by enumerating the outcomes after δt time units, resulting in

$$V\left(d,t\right) \;=\; \max_{r_t}\left\{\, (r_t + V(d+1, t-\delta t)) \cdot \lambda(d, r_t)\delta t + V(d, t-\delta t) \cdot (1 - \lambda(d, r_t)\delta t) + o(\delta t) \right\},$$

where λ is the adoption rate defined in (II.2.3), so $\lambda(d, r_t)\delta t$ is the probability of an adoption if d is the cumulative adoption and r_t is the price. We use this to derive the Hamilton-Jacobi-Bellman (HJB) equation and characterize the first-order condition for the optimal value function. We refer to the optimal pricing policy π^* under a Markovian Bass model as the Markovian Bass pricing (MBP) policy. The following theorem states a relationship between π^* and the value function.

Theorem II.1 (Markovian Bass pricing policy, MBP)**.** Let $r^*(d,t)$ be the price offered under the optimal policy π^* to the Markovian Bass pricing problem (II.3.1) when the $d \in \{0, 1, \ldots, m_0 - 1\}$ is the total past sales and $t \in [0, T]$ is the time remaining in the sales horizon. Then $r^*(d,t)$ is the unique solution to the equation

$$r = -\frac{x(r)}{x'(r)} - V\left(d+1, t\right) + V\left(d, t\right), \tag{II.3.2}$$

where $V(\cdot,\cdot)$ is a function that solves the HJB differential equation

$$\frac{\partial V}{\partial t} + (m_0 - d)\left(p_0 + \frac{d}{m_0} q_0\right) \frac{x(r^*(d,t))^2}{x'(r^*(d,t))} = 0, \tag{II.3.3}$$

with boundary conditions $V(m_0, t) = 0$, for all $t \in [0, T]$, and $V(d, 0) = 0$, for all $d \in \{0, 1, 2, ..., m_0\}$.

The term $\Delta_d V(d,t) := V(d+1,t) - V(d,t)$ that appears in (II.3.2) is the marginal gain in the expected revenue due to an adoption at time t. A myopic seller will choose to maximize the current period expected revenue rate by solving $\max_r \mathbb{E}_{\theta_0}[r \mathrm{d}D_t \mid \mathcal{F}_t]$ in each period. The myopic price satisfies the first order condition $r = -\frac{x(r)}{x'(r)}$. Comparing this condition with (II.3.2), we observe that the sign of $\Delta_d V(d,t)$ informs whether it is optimal to price above or below the myopic seller facing the same conditions. This is the same observation made in the classic paper by *Kalish* (1983) that studies dynamic pricing under the deterministic Bass model (II.2.1). In our notation, *Kalish* (1983) shows (in eq. (9c) of their paper) that the optimal pricing sequence $r^* = (r_t^*, t \geq 0)$ satisfies

$$r_t^* = -\frac{x(r_t^*)}{x'(r_t^*)} - \frac{\mathrm{d}V^{\mathsf{B}}}{\mathrm{d}F_t^r}. \tag{II.3.4}$$

Here, $\mathrm{d}F_t^r$ is the marginal adoption at time t and V^{B} is the optimal expected revenue under the deterministic Bass model. The term $\frac{\mathrm{d}V^{\mathsf{B}}}{\mathrm{d}F_t}$ is referred to as the shadow price $\lambda(t)$ in *Kalish* (1983).

Note the similarity of condition (II.3.4) for the deterministic Bass model to the condition (II.3.2) for the Markovian Bass model. Therefore, the insights from *Kalish* (1983) are also applicable to the dynamic pricing policy under the Markovian Bass model. Specifically, if $\lambda(t) > 0$ or if $\Delta_d V(d,t) > 0$, then there are future benefits of an additional adoption, so the price will be lower than myopic to encourage adoption. Further, if $\lambda(t) < 0$ or if $\Delta_d V(d,t) < 0$, then an additional adoption results in a future loss, so the price will be higher than the myopic price.

In both (II.3.2) and (II.3.4), the optimal price can even be negative if $\lambda(t)$ and $\Delta_d V(d,t)$ are very large. This can occur under a strong imitation effect (i.e., $q_0 \gg p_0$) when the market penetration is still very low (e.g. right after product launch). In this case, temporarily setting a negative price (to encourage fast adoption) is offset by the high value of early adoption. In practice, a negative price can be implemented by seeding early adopters through compensation or perks.[2]

We can also interpret (II.3.2) using the price elasticity as follows. Define the price elasticity of market effort as $e_x := \frac{\mathrm{d}x}{x} \Big/ \frac{\mathrm{d}r}{r}$, where the elasticity evaluated at the optimal price $r = r^*(d,t)$ is e_x^*. From the definition of e_x^*, we have that $\frac{x(r^*)}{x'(r^*)} = \frac{r^*}{e_x^*}$. After substituting this into equation (II.3.2) and rearranging terms, we have

$$r^*(d,t) + \frac{e_x^*}{1+e_x^*} \Delta_d V(d,t) = 0, \tag{II.3.5}$$

where $\frac{e_x^*}{1+e_x^*}$ can be interpreted as the probability of purchasing at price $r^*(d,t)$. Thus, the optimal price $r^*(d,t)$ is the price where the marginal increase in the revenue can offset the expected marginal loss of an adoption.

For the special case of $x(r) = e^{-r}$, we can show that the price elasticity changes

[2] For example, this is a strategy used by the CPG company Johnson & Johnson when it introduces new products (`www.jjfriendsandneighbors.com`). There also exist many influencer programs used by companies such as Fiat, Ford, L'Oreal, or Coca Cola—such as Toluna (`www.toluna.com`), Pinecone Research (`www.pineconeresearch.com`)—which compensate early adopters of products through redeemable points or cash.

proportionally to r, hence the market will not be immediately saturated even if prices are low. For this special case, we are able to derive an analytic expression for the value function V. This special case is interesting since it is a stochastic version of the model considered by *Robinson and Lakhani* (1975). In contrast to *Robinson and Lakhani* (1975), MBP depends on two state variables (instead of one)—the cumulative adoptions and the remaining time until T.

Corollary II.1. If $x(r) = e^{-r}$, then

$$V(d,t) = \ln\left(\sum_{j=1}^{m_0-d} \frac{\prod_{i=d}^{d+j-1} \xi(i)}{j!} \left(\frac{t}{e}\right)^j + 1\right).$$

In general cases, however, the HJB equation cannot be solved analytically. Instead, we can solve (II.3.3) numerically using finite differences, a conventional technique for solving partial differential equations numerically. We describe this method in the e-companion (Appendix A.1).

II.4 Data-driven dynamic pricing with unknown parameters

In this section, we consider the setting where the seller does not know true parameters of the Markovian Bass model, $\theta_0 = (p_0, q_0, m_0)$.

When the true parameter vector θ_0 of the demand model is unknown, one could consider estimating it using historical sales data of like products. This approach is difficult to implement if a like product does not exist, or if the market environment has changed significantly. Other approaches based on subjective expert opinion and on market research are prone to error. One interesting question is: how much revenue does a firm loses when it uses the MBP pricing policy based on wrong parameters? Theorem II.3 later establishes that when wrong model parameters are initially inferred and the data is not used to correct the wrong inference, the revenue loss (relative to R^*) can grow at least linearly in the true market size m_0. This motivates the need to learn the unknown parameters from the price and sales data.

II.4.1 Parameter estimation

Maximum likelihood estimation (MLE) is a method of estimating the unknown θ_0 by choosing the parameters which result in the highest likelihood of observing the data.

The likelihood function is convenient to calculate under the Markovian Bass model. We denote the continuously observed sequence of prices and cumulative sales at time t as

$$\widehat{\mathbf{U}}_t := \left\{ \left(\widehat{r}_s, \widehat{D}_s\right),\ 0 \le s \le t \right\}. \tag{II.4.1}$$

Since the adoption process follows a continuous-time Markov chain, the inter-adoption times are conditionally independent given the previous state information. Let t_i be the time of the ith product adoption, where $i = 0, 1, 2, \ldots$ That is, at time t_k, the cumulative adoption is $\widehat{D}_{t_k} = k$. The likelihood of $\widehat{\mathbf{U}}_t$ under a Markovian Bass model with parameters $\theta = (p, q, m)$ is

$$\ell_t\left(\widehat{\mathbf{U}}_t \mid \theta\right) = \left(\prod_{i=0}^{\widehat{D}_t - 1} \underbrace{\lambda(i, \widehat{r}_{t_{i+1}}; \theta) e^{-\int_{t_i}^{t_{i+1}} \lambda(i, \widehat{r}_s; \theta) ds}}_{f_i(\theta)} \right) \underbrace{e^{-\int_{t_{\widehat{D}_t}}^{t} \lambda(\widehat{D}_t, \widehat{r}_s; \theta) ds}}_{f_{\widehat{D}_t}(\theta)},$$

where $\lambda(i, r; \theta)$ is the instantaneous adoption rate at state i when price is r, which was defined in (II.2.3). Here, $f_i(\theta)$ is the density function of the $(i+1)$-th inter-adoption time, which is mathematically equivalent to the density of inter-arrival times in a non-homogeneous Poisson process with intensity function $\{\lambda(i, \widehat{r}_t; \theta),\ t \ge 0\}$.

Using expression (II.2.3) for $\lambda(i, r; \theta)$, we can rewrite $f_i(\theta)$ as

$$f_i(\theta) := \begin{cases} (m-i)\left(p + \frac{i}{m} q\right) x(\widehat{r}_{t_{i+1}}) \exp\left(-(m-i)\left(p + \frac{i}{m} q\right) \int_{t_i}^{t_{i+1}} x(\widehat{r}_s) \mathrm{d}s \right), & \text{if } i = 0, 1, \ldots, \widehat{D}_t - 1, \\ \exp\left(-(m - \widehat{D}_t)\left(p + \frac{\widehat{D}_t}{m} q\right) \int_{t_{\widehat{D}_t}}^{t} x(\widehat{r}_s) \mathrm{d}s \right), & \text{if } i = \widehat{D}_t. \end{cases} \tag{II.4.2}$$

This results in the following log-likelihood function

$$\begin{aligned} \mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \theta) = & \sum_{i=0}^{\widehat{D}_t - 1} \ln x(\widehat{r}_{t_{i+1}}) + \sum_{i=0}^{\widehat{D}_t - 1} \ln \left[(m-i)\left(p + \frac{i}{m} q\right) \right] - \sum_{i=0}^{\widehat{D}_t - 1} \int_{t_i}^{t_{i+1}} (m-i)\left(p + \frac{i}{m} q\right) x(\widehat{r}_s) \mathrm{d}s \\ & - \int_{t_{\widehat{D}_t}}^{t} (m - \widehat{D}_t)\left(p + \frac{\widehat{D}_t}{m} q\right) x(\widehat{r}_s) \mathrm{d}s. \end{aligned} \tag{II.4.3}$$

The ML estimator $\hat{\theta}_t = (\hat{p}_t, \hat{q}_t, \hat{m}_t)$ maximizes the likelihood of observing the data sequence $\widehat{\mathbf{U}}_t$. That is, $\hat{\theta}_t$ solves the constrained problem $\max_{\theta \ge 0} \mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \theta)$. It is difficult to show the joint concavity of the log-likelihood function in θ. Hence, we perform the following variable transformation:

$$\beta_1 := mp, \quad \beta_2 := q - p, \quad \beta_3 := -\frac{q}{m}, \tag{II.4.4}$$

introduced in *Bass* (1969).[3] We define $\beta_0 := (\beta_{01}, \beta_{02}, \beta_{03})$ to be the transformation variables corresponding to the true Markovian Bass model parameters $\theta_0 = (p_0, q_0, m_0)$.

The log-likelihood function under the transformed variables $\beta = (\beta_1, \beta_2, \beta_3)$ simplifies to:

$$\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \beta) = \sum_{i=0}^{\widehat{D}_t - 1} \ln x(\widehat{r}_{t_{i+1}}) + \int_0^t \ln\left(\beta_1 + \beta_2 \widehat{D}_{s-} + \beta_3 \widehat{D}_{s-}^2\right) \mathrm{d}\widehat{D}_s - \int_0^t \left(\beta_1 + \beta_2 \widehat{D}_s + \beta_3 \widehat{D}_s^2\right) x(\widehat{r}_s) \mathrm{d}s. \tag{II.4.5}$$

The constraint $\theta \geq 0$ implies that $\beta_1 \geq 0$ and $\beta_3 \leq 0$. Hence, the ML estimator $\hat{\beta}_t = (\hat{\beta}_{t1}, \hat{\beta}_{t2}, \hat{\beta}_{t2})$ is the solution to the constrained problem $\max_{\beta:\beta_1 \geq 0, \beta_3 \leq 0} \mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \beta)$. We prove the following proposition that guarantees the tractability this problem.

Proposition II.2. $\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \beta)$ is strictly and jointly concave in β when $\widehat{D}_t \geq 3$.

Proposition II.2 ensures that a standard convex optimization technique such as Newton's method can find the optimizer of $\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \beta)$ efficiently. It also implies identifiability of the ML estimation model of β_0 because the Fisher information matrix is strictly positive definite. This result is useful in establishing the convergence rate of the estimation error (Lemma II.3). Since the log-likelihood function $\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \beta)$ is strictly concave in the transformation variables, it has a unique maximizer, which we denote by $\hat{\beta}_t = (\hat{\beta}_{t1}, \hat{\beta}_{t2}, \hat{\beta}_{t3})$.

If $\hat{\beta}_{t1} > 0$ and $\hat{\beta}_{t3} < 0$, we can readily recover the variables $\hat{\theta}_t = (\hat{p}_t, \hat{q}_t, \hat{m}_t)$ which satisfy the transformation (II.4.4). (Transforming (II.4.2) using (II.4.4), $\hat{\beta}_{t1} = 0$ and $\hat{\beta}_{t3} = 0$ will not happen since the likelihood of this model is zero.) To see how $\hat{\theta}_t$ can be recovered, note that $\hat{m}_t$ solves the equation $\hat{\beta}_{t3}\hat{m}_t^2 + \hat{\beta}_{t2}\hat{m}_t + \hat{\beta}_{t1} = 0$. Since $\hat{\beta}_{t3} < 0$ and $\hat{\beta}_{t1} > 0$, the equation has only one positive root, which we set as $\hat{m}_t$. In this case, $\hat{\theta}_t$ is uniquely determined by the first-order conditions of $\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \beta)$.

An attractive property of ML estimators is that the mean squared error converges to zero with increasing the sample size when the data is independent and identically distributed (*i.i.d.*). Note that, however, inter-adoption times are not identically distributed under the Markovian Bass adoption process. hence the standard argument of ML estimators cannot apply here. *Bradley and Gart* (1962); *Hoadley* (1971) establish the asymptotic properties of ML estimators for independent but not identically distributed samples.

[3]We note that the transformation in *Bass* (1969) was done to perform least squares estimation, and not maximum likelihood.

However, their conditions are difficult to use in our setting. *Roussas* (1969) characterizes regularity conditions to ensure consistency for stationary Markov chains, but the Markovian Bass model is non-stationary. Instead, we follow an approach similar to *Bickel et al.* (2013) and *Broder and Rusmevichientong* (2012) using the concept of Kullback-Leibler divergence from information theory to establish the following Lemma. The lemma characterizes the convergence rate of the mean squared errors of the ML estimators of θ_0.

Lemma II.3. For any fixed time $t > 0$ and $k \geq 3$,

$$\mathbb{E}_{\theta_0}\left((\hat{p}_t - p_0)^2 + (\hat{q}_t - q_0)^2 + \frac{1}{m_0^2}(\hat{m}_t - m_0)^2 \mid D_t^\pi = k\right) \leq \frac{\alpha_\theta}{k+1},$$

where α_θ is a constant that is independent of m_0 and t.

This bound on the estimation error will be crucial in proving a performance bound of the pricing-and-learning algorithms we propose later in this section.

II.4.2 Data-driven pricing policies

Problem (II.2.9) finds a pricing-and-learning policy that maximizes the total expected revenue in the setting where the seller forms an inference about the parameters from the dynamically evolving data. If inference is formed through MLE, then conditional on the information set $\mathcal{F}_t$, the seller infers the demand D_t^π to follow a Markovian Bass model with parameter $\hat{\theta}_t$. In this case, we show in Section A.3 of the appendix that, at any time t, the state variables of the control problem are $(D_t, t, \hat{\theta}_t)$ where D_t is the cumulative sales and $\hat{\theta}_t$ is the current ML estimator. Critical to this state reduction is our result (Proposition A.1) that, conditional on $\mathcal{F}_t$, the future estimate $\hat{\theta}_{t+h}$ is Markovian since the transition probabilities depend only on $(D_t, t, \hat{\theta}_t)$.

Hence, the variables $(D_t, t, \hat{\theta}_t)$ encapsulate the information sufficient to choose a price that maximizes the inferred expected future revenue. This, together with our derived transition probabilities of $\hat{\theta}_{t+h}$, allow us to derive the HJB equation satisfied by the value function. We use this HJB equation to derive the optimal pricing-and-learning policy under ML inference (Theorem A.2).

However, solving for the optimal pricing-and-learning policy suffers from the curse of dimensionality. Specifically, it is more computationally challenging than finding the optimal pricing policy under full information due to additional state variables $\hat{\theta}_t = (\hat{p}_t, \hat{q}_t, \hat{m}_t)$. This motivates us to develop sensible data-driven dynamic pricing policies that (1) utilize data in a computationally efficient way, and (2) have performance guarantees on the

Algorithm II.1: MBP-MLE algorithm

Require: Initial parameters $\hat{\theta}_0 = (\hat{p}_0, \hat{q}_0, \hat{m}_0)$, max horizon length T, subperiod length δ

$s \leftarrow 0$, $\widehat{D}_0 \leftarrow 0$, $\widehat{\mathbf{U}}_{-1} \leftarrow \emptyset$ ▷ *Initialization*

while $s \leq \frac{T}{\delta}$ and $\widehat{D}_s < m_0$ **do**

$\quad r_s \leftarrow \inf \left\{ r : r \geq -\frac{x(r)}{x'(r)} - V(\widehat{D}_s + 1, T - \delta s; \hat{\theta}_s) + V(\widehat{D}_s, T - \delta s; \hat{\theta}_s) \right\}$ ▷ *Set price*

$\quad \widehat{\mathbf{U}}_s \leftarrow \widehat{\mathbf{U}}_{s-1} \cup \{(r_s, \widehat{D}_s + a_s)\}$, where a_s is the new sales in $[\delta s, \delta(s+1))$

$\quad$ **if** $\widehat{D}_s + a_s \geq 3$ **then**

$\quad\quad \hat{\theta}_{s+1} \leftarrow \arg\max_{\theta} \ \mathcal{L}_t \left(\widehat{\mathbf{U}}_s \mid \theta \right)$ ▷ *Update parameter estimate*

$\quad$ **else**

$\quad\quad \hat{\theta}_{s+1} \leftarrow \hat{\theta}_0$

$\quad$ **end if**

$\quad s \leftarrow s + 1$ ▷ *Proceed to next period*

end while

expected revenue. We introduce these policies next, while establishing their performance guarantees in the next section.

The policies utilize the Markovian Bass price function introduced in Section II.3 where they replace the true (unknown) parameters $\theta_0 = (p_0, m_0, q_0)$ with parameter estimates. Therefore, with slight abuse of notation, we define $r_t^*(\theta, d)$ as the Markovian Bass price function (Theorem II.1) if t is the elapsed time since introducing the product, d is the number of past adoptions, and θ is the demand parameter vector.

II.4.2.1 Pricing policy with continuous price changes.

We first propose a policy referred to as MBP-MLE (outlined in Algorithm II.1). At time t, this policy offers the Markovian Bass price of Section II.3 except that when computing the price and the value function, it replaces the true (unknown) parameter vector θ_0 with the ML estimator $\hat{\theta}_t$. Note that the Markovian Bass price is optimal if there is no error in parameter estimation. Hence, the price offered by the MBP-MLE policy exploits the current estimate and ignores the role of price in improving future inference. Yet, our analysis in the next section shows that the revenue loss of MBP-MLE grows sublinearly in the market size at the rate of $\mathcal{O}(\ln m_0)$ (Theorem II.4).

The MBP-MLE policy starts with an initial guess of the parameters, $\hat{\theta}_0 = (\hat{p}_0, \hat{q}_0, \hat{m}_0)$, which it uses to set the prices $r_t = r_t^*(\hat{\theta}_0, \widehat{D}_t)$, for any t where $\widehat{D}_t < 3$. The quality of the initial guess does not affect the policy's performance bound. As time progresses, more observations are added to the data $\widehat{\mathbf{U}}_t := \{(\widehat{r}_s, \widehat{D}_s) \mid s \leq t\}$. When $\widehat{D}_t \geq 3$, the seller can use the accrued data to solve for the ML solution $\hat{\theta}_t$ and set the price at $r_t = r_t^*(\hat{\theta}_t, \widehat{D}_t)$.

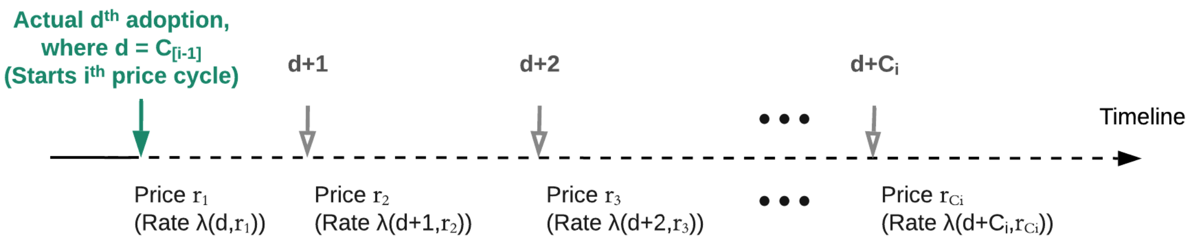


Figure II.2: Certainty equivalent MBP-MLE prices and adoption rates in one price cycle. The green arrow indicates when the cycle starts. The gray arrows are the deterministic adoption times.

II.4.2.2 Pricing policy with limited price changes.

In many situations, frequent price changes can be difficult or impractical to implement due to cost, time and loss of goodwill associated with price changes. This explains why many firms only change price a few times during the season.

We next propose the MBP-MLE-Limited policy in which the firm changes its price at most K times. One way to model this is to include the number of price changes as a state variable. However, doing so will further increase the complexity of the dynamic programming model. Instead, we propose a simpler approach by assuming that price changes occur when the cumulative adoption reaches certain thresholds (e.g, the 100th customer, the 1000th customer, etc). This approach of using cumulative purchases as triggers for price changes has been used in selling new products by the crowdfunding platforms KickStarter and IndieGoGo (*Stonemaier Games*, 2013).

Consider a sequence of natural numbers $C := \{C_i,\ i = 0, 1, 2, \ldots, K\}$, where $C_i \geq 1$ for any i. Define $C_{[-1]} := 0$ and $C_{[i]} := \sum_{k=0}^{i} C_k$ for all $i = 0, \ldots, K$. For the ith price cycle, our proposed MBP-MLE-Limited policy sets the same price $r^{(i)}$ starting from when the $C_{[i-1]}$-th adoption has occurred until when the $C_{[i]}$-th adoption happens. Hence, unless the end of the horizon is reached, the per-unit revenue $r^{(i)}$ will be earned by the seller from exactly C_i adopters. For now, we will assume that K and C are both given. Later in Section II.5.3, we describe how K and C can be chosen, even without knowing m_0, so that the revenue loss of MBP-MLE-Limited is $\mathcal{O}(\ln m_0)$ (Theorem II.5).

We next describe how the policy determines the prices for each cycle. Suppose that the ith price cycle has just been triggered at time t by the adoption of the $C_{[i-1]}$-th customer. After updating the ML estimator $\hat{\theta}$, MBP-MLE-Limited chooses a price $r^{(i)}$ for the next C_i adoptions. The idea is that the total revenue under $r^{(i)}$ is set to match the expected revenue if MBP-MLE could be used in the upcoming cycle (i.e., where each of the C_i customers is charged a different price). Note that there are no actual price changes during the cycle after the initial price change. The MBP-MLE prices are only used to construct the lookahead value to compute $r^{(i)}$.

The lookahead value is constructed using the certainty equivalent of the MBP-MLE prices and the corresponding adoption rates for the ith cycle (see Figure II.2). In the figure, vertical arrows correspond to times of adoptions. The green solid arrow is the actual $C_{[i-1]}$-th adoption that triggers a price cycle. The gray, empty arrows are the predicted future adoption times using a deterministic model. We denote $d = C_{[i-1]}$ for notational convenience. As illustrated in the figure, once the MBP-MLE price r_j is set (i.e., immediately after the $(d + j - 1)^{\text{th}}$ customer purchases where $j = 1, \dots, C_i$), the adoption rate changes to $\lambda(d + j - 1, r_j; \hat{\theta})$ where λ is defined in (II.2.3). Hence, the expected inter-adoption time between the $(d + j - 1)^{\text{th}}$ and $(d + j)^{\text{th}}$ adoption is

$$\Delta t_j := \frac{1}{\lambda(d + j - 1, r_j; \hat{\theta})} = \frac{1}{\xi(d + j - 1; \hat{\theta})x(r_j)}. \tag{II.4.6}$$

This then determines the time of the next adoption $d + j$ under a deterministic model, assuming that the previous inter-adoption times $\Delta t_1, \Delta t_2, \dots, \Delta t_{j-1}$ have already been computed. Since the MBP-MLE prices depend only on the elapsed time $\tau = t + \sum_{k=1}^{j} \Delta t_k$ and the cumulative adoptions $d + j$, this allows us to compute the next price $r_{j+1} := r^*_\tau(\hat{\theta}, d + j)$, which determines the next inter-adoption time Δt_{j+1}. This proceeds until we have the complete deterministic sequence of MBP-MLE prices for the ith cycle.

Given the MBP-MLE price sequence $\{r_1, \dots, r_{C_i}\}$ for the ith cycle, the MBP-MLE-Limited policy then chooses the price $r^{(i)}$ to satisfy the following relation:

$$\sum_{j=1}^{C_i} r^{(i)} \lambda \left(d + j - 1, r^{(i)}; \hat{\theta} \right) \Delta t_j = \sum_{j=1}^{C_i} r_j \lambda \left(d + j - 1, r_j; \hat{\theta} \right) \Delta t_j.$$

The right-hand side is the certainty equivalent revenue of MBP-MLE in the ith cycle. Hence, the MBP-MLE-Limited price is chosen such that its expected revenue matches the certainty equivalent revenue of MBP-MLE, assuming that the adoption times of the C_i customers are fixed. From identity (II.4.6), $r^{(i)}$ is the solution to

$$r^{(i)} x \left(r^{(i)} \right) = \frac{\sum_{j=1}^{C_i} r_j}{\sum_{j=1}^{C_i} 1/x(r_j)}. \tag{II.4.7}$$

For the first price cycle ($i = 0$), we assume that the policy starts with an initial guess of parameters $\hat{\theta}_0$. We also assume that $C_0 \geq 3$ so that there exists an ML estimator when the first price change is calculated. Algorithm II.2 provides the outline for the

MBP-MLE-Limited algorithm.

Algorithm II.2: MBP-MLE-Limited algorithm

Require: Initial parameters $\hat{\theta}_0 = (\hat{p}_0, \hat{q}_0, \hat{m}_0)$, max horizon length T, subperiod length δ, and price change triggers $\{C_i, i = 0, 1, 2, ..., K\}$ where $C_0 \geq 3$

function LIMITED-PRICE(θ, C, d, t)
  for $j \leftarrow 1, 2, \ldots, C$ **do**
    $dV \leftarrow V(d + j, t; \theta) - V(d + j - 1, t; \theta)$
    $r_j \leftarrow \inf\left\{r : r \geq -\frac{x(r)}{x'(r)} - dV\right\}$ ▷ *Calculate MBP-MLE for adoption $d + j - 1$*
    $\Delta t_j \leftarrow \frac{1}{\xi_{d+j-1}(\theta) x(r_j)}$ ▷ *Approximate the inter-adoption time for $d + j$*
    $t \leftarrow t - \Delta t_j$
  end for
  $\bar{r} \leftarrow \inf\left\{r : r \cdot x(r) \geq \frac{\sum_{j=1}^{C} r_j}{\sum_{j=1}^{C} 1/x(r_j)}\right\}$ ▷ *Calculate the price for the C adoptions*
  return $\bar{r}$
end function

$r_0 \leftarrow$ LIMITED-PRICE$(\hat{\theta}_0, C_0, 0, T)$, $s \leftarrow 1$, $i \leftarrow 1$, $\widehat{D}_0 \leftarrow 0$, $\widehat{\mathbf{U}}_0 \leftarrow \emptyset$ ▷ *Initialization*
while $s \leq \frac{T}{\delta}$ and $\widehat{D}_s < m_0$ **do**
  $\widehat{\mathbf{U}}_s \leftarrow \widehat{\mathbf{U}}_{s-1} \cup \{(r_s, \widehat{D}_s + a_s)\}$, where a_s is the new sales in $[\delta(s-1), \delta s)$ ▷ *Update dataset*
  if $a_s = 1$ **and** $\widehat{D}_s + a_s = \sum_{k=0}^{i-1} C_k$ **then** ▷ *Price change is triggered*
    $\hat{\theta}_s \leftarrow \arg\max_{\theta} \mathcal{L}_t\left(\widehat{\mathbf{U}}_s \mid \theta\right)$ ▷ *Update parameter estimate*
    $r_s \leftarrow$ LIMITED-PRICE$(\hat{\theta}_s, C_i, \widehat{D}_s, T - \delta s)$ ▷ *Change price*
    $i \leftarrow i + 1$ ▷ *Increase number of price changes*
  else
    $r_s \leftarrow r_{s-1}$ ▷ *Do not change price*
  end if
  $s \leftarrow s + 1$ ▷ *Proceed to next period*
end while

II.5 Analysis of pricing policies

We next characterize the performance of our proposed pricing-and-learning policies, MBP-MLE and MBP-MLE-Limited. We do this by deriving analytic bounds on the gap of their expected revenues against the oracle revenue R^* defined in (II.3.1). Since R^* is the expected revenue of the oracle policy π^* that knows θ_0, then the expected revenue of any pricing-and-learning policy cannot exceed R^*. Hence, for any pricing-and-learning policy π with expected revenue $R(\pi)$, the difference $R^* - R(\pi)$ can be viewed as the revenue loss of π.

In this section, we derive asymptotic bounds on the revenue loss of our proposed policies as the market size m_0 grows. We will establish that the revenue losses of our proposed policies grow at most sublinearly with m_0 at the rate $\mathcal{O}(\ln m_0)$.

The challenge in bounding the revenue loss when demand follows a Markovian Bass model is that pricing mistakes affect, not only the current revenue, but also the revenues in any future time period. This is because adoption rates (hence, revenues) depend on the cumulative adoptions, which in turn can be influenced by prices from any past period. Hence, the effects of pricing mistakes can compound over time. To bound the revenue loss, we then need to establish a non-stationary relationship among revenue loss, pricing errors and estimation errors.

Let $D^* = (D_t^*,\ t \geq 0)$ and $r^* = (r_t^*,\ t \geq 0)$ denote the cumulative adoption process and the price process, respectively, under the oracle policy π^*. The following proposition states a general result for any pricing policy in the set of $\mathcal{F}_t$-adapted policies Π. The definitions of $\mathcal{O}, \Omega$, and Θ can be found in Section II.1.2. Particularly, we are interested in the limiting behavior as m_0 goes to infinity.

Proposition II.3. If $\pi \in \Pi$ and there exists an $\alpha > 0$ independent of t and m_0 such that $\mathbb{E}_{\theta_0}[|r_t^\pi - r_t^*|] \geq \alpha t e^{-t} {m_0}^{-1}$ for $t \in [0, T]$, then

$$R^* - R(\pi) = \mathcal{O}\left(\mathbb{E}_{\theta_0}\left[\int_0^T \frac{D_t^\pi + 1}{t + t_0}(r_t^\pi - r_t^*)^2 \mathrm{d}t\right]\right), \tag{II.5.1}$$

where $t_0 = \Theta({m_0}^{-1})$.

This important proposition establishes that the revenue loss of a policy π is bounded by a weighted average of its squared pricing errors relative to the oracle policy π^*. Since any past pricing error can linger and affect future adoptions, the weights represent the cumulative effect of the pricing error on the revenue loss.

The idea behind the proof of Proposition II.3 is that we can decompose the revenue loss into two parts:

$$R^* - R(\pi) \leq \mathbb{E}_{\theta_0}\left[\int_0^T |x(r_t^*)r_t^* - x(r_t^\pi)r_t^\pi| \cdot \xi(D_t^*)\mathrm{d}t\right] + \mathbb{E}_{\theta_0}\left[\int_0^T x(r_t^\pi)r_t^\pi \cdot |\xi(D_t^\pi) - \xi(D_t^*)|\, \mathrm{d}t\right]$$

Note that at time t, the firm accrues revenue at the rate $r_t^\pi \xi(D_t^\pi) x(r_t^\pi)$, where the adoption rate exhibits a price effect $x(r_t^\pi)$ and a word-of-mouth effect $\xi(D_t^\pi)$. The first term in the decomposition measures the revenue loss in the current period only since it assumes that the word-of-mouth effect is the same as the oracle policy. The second term captures the revenue loss due to the compounded word-of-mouth effects. We prove that, under Assumption II.1, the first part grows in the order of the squared price difference. As the mean difference in the proportions of adopters under π and π^* vanishes at a fast rate (Lemma A.2), the second part is dominated by the first part. The full proof can be found in the appendix.

The implication of Proposition II.3 is that, to bound the revenue loss of a policy π, it suffices to bound the squared price error between π and π^*. Our next result is important since it establishes such a price error bound for policies that use the Markovian Bass pricing function with a parameter sequence $\{\theta_t, t \geq 0\}$ (i.e., at time t, offer the price $r_t^*(\theta_t, D_t)$ where D_t is the cumulative adoption). Note that both MBP-MLE and MBP-MLE-Limited are policies of this type.

Lemma II.4. Consider a parameter sequence $\{\theta_t = (p_t, q_t, m_t), t \geq 0\}$ where θ_t is an $\mathcal{F}_t$-measurable random vector, and p_t, q_t, m_t are finite, $p_t + q_t > 0$ and $m_t > 0$ for all $t \geq 0$ almost surely. If π is the policy that offers the Markovian Bass price with parameter θ_t, i.e., $r_t^\pi = r_t^*(\theta_t, D_t^\pi)$, then for any $t \in (0, T]$,

$$\mathbb{E}_{\theta_0}\left[(r_t^* - r_t^\pi)^2 \mid \mathcal{F}_t\right] = \Theta\left(\mathbb{E}_{\theta_0}\left[(\hat{p}_t - p_0)^2 + (\hat{q}_t - q_0)^2 + \frac{1}{m_0^2}(\hat{m}_t - m_0)^2 \mid \mathcal{F}_t\right]\right) + \mathcal{O}\left(\frac{1}{m_0}\right). \tag{II.5.2}$$

The proof is in the appendix. The proof decomposes the squared pricing error as follows:

$$\mathbb{E}_{\theta_0}\left[(r_t^* - r_t^\pi)^2\right] = \Theta\left(\mathbb{E}_{\theta_0}\left[(r_t^*(\theta_0, D_t^*) - r_t^*(\theta_t, D_t^*))^2\right]\right) + \Theta\left(\mathbb{E}_{\theta_0}\left[(r_t^*(\theta_t, D_t^*) - r_t^*(\theta_t, D_t^\pi))^2\right]\right)$$

where both terms on the right-hand side are potentially affected by the market size m_0. The first term in the decomposition represents the pricing error originating from the parameter estimation error, and the second term represents the pricing error originating from the difference in cumulative adoptions (D_t^* and D_t^π) as a result of price differences up to time t. This second term reflects the fact that the deviation of r_t^π from r_t^* is not only from estimation errors but also from differences in the cumulative adoptions.

Lemma II.4 and Proposition II.3 together enable us to bound the revenue loss by the parameter estimation error. Hence, the revenue loss of a data-driven pricing policy π can be analyzed by studying the dynamics of parameter estimation errors under π.

Before proceeding with our analysis, we first derive a fundamental limit on the revenue loss of data-driven pricing policies. This is usually accomplished by constructing a special case of the problem that satisfies Assumption II.1, and showing that for any data-driven pricing policy, the worst-case revenue loss associated with that special case cannot be lower than the fundamental limit (see *Broder and Rusmevichientong* 2012; *Besbes et al.* 2015). The following theorem states that $\Omega(\ln m_0)$ is the fundamental limit under the special case of $x(r) = e^{-r}$. (Note that this functional form satisfies Assumption II.1.) Hence, the fundamental limit $\Omega(\ln m_0)$ serves as a benchmark for the revenue loss of data-driven pricing policies in our problem setting.

Theorem II.2. Let $x(r) = e^{-r}$ with $r \in [0, 2)$. Then for any pricing-and-learning policy $\pi \in \Pi$, there exists a true value $q_0 \in [1/4, 5/4]$ such that $R^* - R(\pi) = \Omega(\ln m_0)$.

The proof of Theorem II.2 requires showing that the pricing error is lower-bounded by a rate inversely proportional to the sample size. This is formalized in Claim A.3 using the Bayesian Cramer-Rao inequality (also known as van Trees' inequality, see Lemma A.4), which provides a lower bound on the performance of sequential decision policies. We then use a tight version of Proposition II.3 (Claim A.2) to connect the pricing error to revenue loss.

In the remainder of this section, we will analyze the expected revenue loss of several pricing-and-learning policies, including our proposed policies MBP-MLE and MBP-MLE-Limited. In the analyses, we assume $x(\cdot)$ is any function that satisfies Assumption II.1. In fact, we will show that the revenue losses of our proposed policies are $\mathcal{O}(\ln m_0)$.

II.5.1 Revenue loss without learning

Consider a pricing policy π^s that offers the Markovian Bass prices based on an initial estimate $\hat{\theta}_0$ that is never updated even when data is available. In fact, relying on an initial estimate is the approach suggested in many papers including *Bass* (1969); *Bass et al.* (1994) and *Krishnan et al.* (1999). Theorem II.3 below states that the revenue loss of such a policy can be large and grows at least linearly in m_0. Establishing the result requires the following additional condition on $x(\cdot)$.

Assumption II.2. The marketing effort function $x : \mathbb{R} \to \mathbb{R}^+$ has the property that there exists a constant $\underline{C} > 0$ such that $\left|\frac{\partial^2}{\partial r^2}(rx(r))\right| \geq \underline{C}$ for all $-\infty < r < \infty$.

Assumption II.2 is not restrictive since it is easily satisfied as long as the instantaneous revenue rate, $rx(r)$, is strictly concave in r, a standard assumption in the revenue management literature.

Theorem II.3. Given a parameter estimate $\hat{\theta}_0$, let π^s be the pricing policy that offers the price $r_t^*(\hat{\theta}_0, D_t^s)$ at time t where D_t^s is the cumulative adoptions by time t. Under Assumption II.2 and if $\|\theta_0 - \hat{\theta}_0\|^2 = \mathcal{E}^2$, then $R^* - R(\pi^s) = \Omega(\mathcal{E}^2 m_0)$.

II.5.2 Revenue loss of MBP-MLE

We next establish an upper bound on the revenue loss of MBP-MLE. Unlike the simple pricing policy π^s, the MBP-MLE policy π^{M} continuously updates the parameters of the Markovian Bass price function using the ML estimators.

The implication from Proposition II.3 and Lemma II.4 is that the revenue loss of any pricing-and-learning policy that uses the Markovian Bass price function with parameter estimates depends only on the weighted mean squared error of those parameter estimates. We can therefore use our bound on the estimation error of MLE in a Markovian Bass model (Lemma II.3) to obtain a performance guarantee for the MBP-MLE policy. This is formally stated in the following theorem. The detailed proof is provided in the e-companion.

Theorem II.4. If π^{M} is the MBP-MLE policy, then $R^* - R(\pi^{\mathsf{M}}) = \mathcal{O}(\ln m_0)$.

Note that Lemma II.3, Lemma II.4, and Proposition II.3 are important results for establishing the upper bound. Intuitively, we have a $\mathcal{O}(\ln m_0)$ bound since the estimation error at time t is inversely proportional to $D_t + 1$ (Lemma II.3), which incidentally is also the weight applied to the pricing error in (II.5.1). The detailed proof of the theorem is in the appendix.

Note that MBP-MLE fully exploits the current parameter estimate since the resulting MBP price is not adjusted to improve the accuracy of parameter estimation. Despite not actively doing price exploration, the revenue loss of MBP-MLE, $\mathcal{O}(\ln m_0)$, matches the fundamental lower bound on the revenue loss of any data-driven pricing policy (Theorem II.2). In Section II.5.5, we will discuss why learning appears to occur for free under MBP-MLE.

II.5.3 Revenue loss of MBP-MLE-Limited

We now derive a performance bound for the MBP-MLE-Limited policy $\pi^{\mathsf{M\text{-}Lim}}$, a policy with limited price changes. Recall that this policy requires a sequence $\{C_0, C_1, \ldots, C_K\}$ to determine the number of adoptions between price changes. In our asymptotic analysis where the potential market size m_0 grows, it is reasonable that either the number of price changes, K, increases or the number of adoptions between price changes increases. In either case, we assume that $\sum_{i=0}^{K} C_i = \Theta(m_0)$.

The following result establishes an asymptotic bound on the revenue loss of MBP-MLE-Limited.

Theorem II.5. Let $\pi^{\textsf{M-Lim}}$ be the MBP-MLE-Limited pricing policy where $\{C_0, C_1, \ldots, C_K\}$ are the number of adoptions between price changes with $C_0 \geq 3$. Then,

$$R^* - R\left(\pi^{\textsf{M-Lim}}\right) = \mathcal{O}\left(\max\left\{C_0,\, 1 + \max_{i=1,2,\ldots,K} \frac{C_i}{C_{i-1}}\right\} \cdot \ln m_0\right).$$

Compared with MBP-MLE, the cumulative pricing error originating from inaccurate parameter estimates is larger, since the estimates are only updated at the price change points. However, if the firm chooses price change points such that the number of adoptions between price changes grows exponentially large, the revenue loss grows at most in logarithmic order. To see this, note that if $C_0 = 3$ and $C_i = C_0 a^i$ for all $i = 1, \ldots, K$ for some base $a \geq 2$, then $R^* - R\left(\pi^{\textsf{M-Lim}}\right) = \mathcal{O}\left((1+a)\ln m_0\right)$. Hence, the revenue loss bound still remains in the same order as that of MBP-MLE. However, the number of price changes is $\Theta(\ln m_0)$, while MBP-MLE implements continuous price changes.

When $C_i = C_0 a^i$, most price changes occur during the early stages of adoption so that the firm can collect enough information. At the later stages of adoption, the firm simply exploits this and uses a relatively stable pricing strategy. Hence, price experimentation primarily occurs at the start of the launch. Doing so can prevent significant revenue loss overall. Note that using an exponentially growing sequence for C_i resembles many learning-while-doing policies in the literature (e.g., *Cheung et al.* 2017 and *Qi et al.* 2017).

On the other hand, the firm can have a revenue loss that grows superlinearly when it chooses a non-increasing sequence $\{C_i, i = 0, 1, 2, \ldots, K\}$, such as a decreasing or constant sequence. For example, if $C_0 = C_1 = \ldots = C_K = \Theta(m_0)$, then the expected revenue loss can be as large as $\Theta\left(m_0 \ln m_0\right)$. If the adoptions between price changes is non-increasing over time, then to achieve $\mathcal{O}(\ln m_0)$ growth, the number of price changes must be sufficiently large (at least in the order of m_0). The reason is, with a non-increasing sequence, since $\sum_{i=0}^{K} C_i = \Theta(m_0)$, this implies $(K+1)C_0 \geq \Theta(m_0)$. To maintain the revenue loss to be bounded by $\ln m_0$, we need $C_0 = O(1)$.

Finally, we also comment on our choice of using adoption numbers to trigger price changes. With this method, the number of adoptions are known when the ML estimates are updated, so we can utilize our previous result on ML estimation errors (Lemma II.3) in proving the bound on revenue loss. One could also consider a pricing policy where a price changes are triggered by time (e.g., every Monday at 8 a.m.). While the two policies do not differ much in terms of execution, characterizing the estimation accuracy in the latter is harder to do because the cumulative number of adoptions at each price change period is a random variable.

II.5.4 Extension to an unknown marketing effort function

The two algorithms (MBP-MLE and MBP-MLE-Limited) and their asymptotic analysis can be extended to the case where the marketing effort function $x(\cdot)$ is unknown. This can be done if x is a Bernstein polynomial with unknown parameters.

Let $x(r;\gamma) = \sum_{i=0}^{n} \gamma_i b_{i,n}(r)$ where n is the order of the polynomial, $b_{i,n}(r) = \binom{n}{i} r^i (1-r)^{n-i}$ are the Bernstein basis functions, and $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_n) \in \mathbb{R}^{n+1}$ is a parameter vector. (This requires price to be normalized to $[0,1]$. This can be done without loss of generality for any price r defined on $[\underline{r}, \overline{r}]$ since we can transform our model by introducing a new variable $(r-\underline{r})/(\overline{r}-\underline{r})$.) Bernstein polynomials are known to be able to approximate any continuous function defined on $[0,1]$ (*Lorentz*, 2013). It has been proven that the Bernstein polynomial approximation converges to the true function uniformly at a rate of $n^{-1/2}$ (*Lorentz*, 2013). Thus, assuming the marketing effort function is a Bernstein polynomial is quite general. For example, the commonly used (see *Robinson and Lakhani* 1975 and *Chow* 1960) market effort function $x(r) = a - br$ can be considered as a Bernstein polynomial and $x(r) = e^{a-br}$ can be well approximated by Bernstein polynomials.

We assume the seller knows that $x(\cdot)$ is a Bernstein polynomial of order n, but she does not know the true parameter vector, which we denote as $\gamma_0 = (1, \gamma_{0,1}, \gamma_{0,2} \ldots, \gamma_{0,n})$. The seller uses maximum likelihood to estimate the $n+4$ Markovian Bass model parameters $(p_0, q_0, m_0, \gamma_0)$. Note that we normalized the vector γ_0 such that $\gamma_{0,0} = 1$. This can be done without loss of generality, and we will later show that this makes the model identifiable under ML estimation.

Let $\mu = (\beta, \gamma)$, where β is the transformation defined in (II.4.4). The function $\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \mu)$ is not necessarily jointly concave in μ. However, we will show that after a proper transformation of the parameters, the log-likelihood function is strictly and jointly concave in the transformed parameters when the data has an initial price exploration and there is a sufficient number of adoptions. Specifically, consider the following transformation:

$$\mu' := (\gamma_j \beta_1,\ \gamma_j \beta_2,\ \gamma_j \beta_3, \quad j = 0, 1, \ldots, n)^\top. \tag{II.5.3}$$

Here, μ' is a vector of size $3(n+1)$. By definition, $\mu'_{3j+\ell} = \gamma_j \beta_\ell$ where $j = 0, 1, \ldots, n$ and $\ell = 1, 2, 3$. At time t, for each adoption $i = 0, 1, \ldots, \widehat{D}_t$, we can construct the following $3(n+1)$-dimensional column vectors from the data $\widehat{\mathbf{U}}_t$:

$$y^{i,s} := \left(b_{j,n}(\widehat{r}_s),\ b_{j,n}(\widehat{r}_s) \cdot i,\ b_{j,n}(\widehat{r}_s) \cdot i^2, \quad j = 0, 1, \ldots, n\right)^\top, \qquad \text{for any } s \in [t_i, t_{i+1}]$$

Then, the log-likelihood function under the transformed parameters μ' simplifies to:

$$\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \mu') = \sum_{i=0}^{\widehat{D}_t-1} \ln \mu'^{\top} y^{i,t_{i+1}} - \sum_{i=0}^{\widehat{D}_t-1} \int_{t_i}^{t_{i+1}} \mu'^{\top} y^{i,s} \mathrm{d}s - \int_{t_{\widehat{D}_t}}^{t} \mu'^{\top} y^{\widehat{D}_t,s} \mathrm{d}s. \tag{II.5.4}$$

It is easy to check that $\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \mu')$ is jointly concave in μ'. In fact, Proposition II.4 next states that it is strictly concave under some condition on the initial prices.

Proposition II.4. If $\widehat{D}_t \geq 3(n+1)$ and if the price sequence $(\widehat{r}_{t_{i+1}}, i = 0, \ldots, 3n+2)$ is chosen such that the matrix

$$\mathbf{Y} := \begin{pmatrix} y^{0,t_1} & y^{1,t_2} & \cdots & y^{3n+2,t_{3n+3}} \end{pmatrix} \in \mathbb{R}^{3(n+1)\times 3(n+1)} \tag{II.5.5}$$

has full rank, then the log-likelihood function $\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \mu')$ is strictly and jointly concave in μ'.

Note that the condition that $\mathbf{Y}$ is full rank is a condition on price exploration. Intuitively, this condition can be achieved if the prices offered to the first $3(n+1)$ adoptions are sufficiently different. Hence, μ' is identifiable under MLE if there is an initial price exploration phase.

We next discuss how to recover μ from μ'. Due to our normalization $\gamma_0 = 1$, we have $\beta_1 = \mu'_1$, $\beta_2 = \mu'_2$, and $\beta_3 = \mu'_3$. Furthermore, for any $j = 1, \ldots, n$, we have $\gamma_j = \mu'_{3j+1}/\mu'_1 = \mu'_{3j+2}/\mu'_2 = \mu'_{3j+3}/\mu'_3$. Given the data $\widehat{\mathbf{U}}_t$, we can find the ML estimator of μ' by solving:

$$\begin{aligned} \max_{\mu'} \ & \mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \mu') \\ \text{s.t. } & \mu'_1 \geq 0, \ \mu'_3 \leq 0 \\ & \mu'_{3j+1}/\mu'_1 = \mu'_{3j+2}/\mu'_2, \quad j = 1, \ldots, n \\ & \mu'_{3j+2}/\mu'_2 = \mu'_{3j+3}/\mu'_3, \quad j = 1, \ldots, n \end{aligned} \tag{II.5.6}$$

We denote the solution to (II.5.6) as $\hat{\mu}'_t$. We can then construct $\hat{\mu}_t = (\hat{\beta}_t, \hat{\gamma}_t)$ from the solution $\hat{\mu}'_t$.

Optimization model (II.5.6) has a strictly concave objective function (under the Proposition II.4 condition) and a non-convex feasible set (due to nonlinear equality constraints). Hence, we cannot use efficient techniques for convex optimization. Observe, however, that if (μ'_1, μ'_2, μ'_3) is fixed, then the problem has linear equality constraints, so the feasible set is convex. Therefore, a method for solving (II.5.6) is to search for the largest log-likelihood value over the space (μ'_1, μ'_2, μ'_3) where, at each point in the space, a strictly concave function is maximized subject to linear equality constraints.

We next adapt the data-driven pricing policies MBP-MLE and MBP-MLE-Limited to the case when $x(\cdot)$ is an unknown Bernstein polynomial. Whenever a parameter estimate is required in MBP-MLE and MBP-MLE-Limited, we use the ML estimators of $(p_0, q_0, m_0, \gamma_0)$. In fact, we can establish the rate of convergence of ML estimators, similar to Lemma II.3. This is because the log-likelihood function $\mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \mu)$ is continuously differentiable and element-wise concave in all parameters. As a result, all arguments in the proof of Lemma II.3 apply to ML estimators of $\mu_0 = (\beta_0, \gamma_0)$. This gives us the following Lemma II.3′.

Lemma II.3′. For any fixed time t, if $k \geq 3(n+1)$ and $\mathbf{Y}$ defined in (II.5.5) is full rank, then

$$\mathbb{E}_{\mu_0}\left(\|\hat{\beta}_t - \beta_0\|^2 \mid D_t^\pi = k\right) \leq \frac{\alpha_\beta}{k+1}, \quad \text{and}$$
$$\mathbb{E}_{\mu_0}\left(\|\hat{\gamma}_t - \gamma_0\|^2 \mid D_t^\pi = k\right) \leq \frac{\alpha_\gamma}{k+1},$$

where $\alpha_\beta, \alpha_\gamma$ are constants that are independent of m_0.

The next step is to establish the relationship between the pricing errors and the estimation errors, similar to Lemma II.4. Let $r_t^*(\mu, d)$ be the Markovian Bass price for parameters $\mu = (\beta, \gamma)$ when the cumulative adoption is d and the elapsed time is t. Below is our next result.

Lemma II.4′. Consider a parameter sequence $\{\mu_t = (p_t, q_t, m_t, \gamma_{t,0}, \ldots, \gamma_{t,n}),\ t \geq 0\}$ where μ_t is an $\mathcal{F}_t$-measurable random vector, and $p_t, q_t, m_t, \gamma_{t,j}$ are finite, and $p_t + q_t > 0$ and $m_t > 0$ for all $t \geq 0$ and $j = 0, \ldots, n$ almost surely. If π is the policy that offers the Markovian Bass price with parameter μ_t, i.e., $r_t^\pi = r_t^*(\mu_t, D_t^\pi)$, then for any $t \in (0, T]$,

$$\mathbb{E}_{\mu_0}\left[|r_t^* - r_t^\pi|^2 \mid \mathcal{F}_t\right] = \Theta\left(\mathbb{E}_{\mu_0}\left[\|\mu_t - \mu_0\|^2 \mid \mathcal{F}_t\right]\right) + \mathcal{O}\left(\frac{1}{m_0}\right)$$

for some $\alpha > 0$ independent of m_0.

Therefore, utilizing Lemma II.3′, Lemma II.4′ and Proposition II.3, we derive similar results as Theorem II.4 and Theorem II.5 under the extension to an unknown marketing effort function.

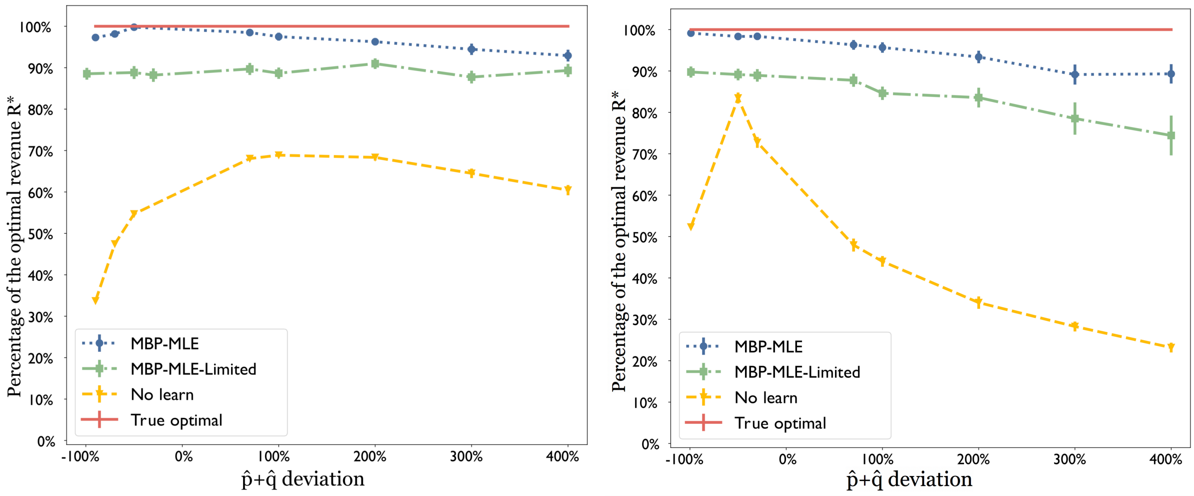
II.5.5 Discussion of why learning occurs for “free"

It should be noted that our bound $\mathcal{O}(\ln m_0)$ on MBP-MLE and MBP-MLE-Limited coincides with the fundamental lower bound on revenue loss in Theorem II.2. It also coincides with the lower bound derived in *Broder and Rusmevichientong* (2012) for the class of well-separated problems. The *well-separated* condition means that any two distinct parameters would generate non-intersecting expected demand curves. Although the model in our setting is past dependent, the $\mathcal{O}(\ln m_0)$ revenue loss of MBP-MLE and MBP-MLE-Limited still matches the fundamental lower bound.

When $x(\cdot)$ is known, it is surprising that the lower bound is achieved even if both policies do not explicitly change price for the purpose of increasing learning accuracy (i.e., experiment with price). In fact, both policies exploit the current information by using the ML estimates as if they are the true parameters. We call this “learning-for-free.” However, “learning-for-free” does not always happen when $x(\cdot)$ is unknown. As shown in Proposition II.4, we need an initial price exploration phase to ensure the model parameters can be uniquely identified. But, free learning occurs after this initial price exploration phase.

We next discuss why learning occurs for free when $x(\cdot)$ is known and the parameters of the Markovian Bass model are estimated using maximum likelihood. Note that the log-likelihood function (II.4.3) is changing continuously over time even when price is unchanged. Hence, the ML estimators are continuously updated in time regardless of the price path. From Lemma II.3, the accuracy of the ML estimators increases as more people adopt. Indeed, the ML estimators will converge to the true parameters under any pricing policy as time t increases. The parameters will also converge under pricing that exploits the current parameter estimates. Hence, exploration and exploitation can occur simultaneously when the parameters are estimated using the maximum likelihood.

MBP-MLE allows continuous price changes, so the benefit from the increasingly accurate ML estimators is immediately realized through pricing that exploits the current estimates. This explains why the revenue loss is $\mathcal{O}(\ln m_0)$ even without changing prices for the explicit purpose of price exploration. On the other hand, MBP-MLE-Limited has limited opportunities to change prices. Though the ML estimators are continuously updated and will converge to the true parameters with more adoptions (c.f. Lemma II.3 and Lemma II.3′), the changes in the estimators are only reflected onto the price at the limited price change epochs. Therefore, a key to limiting the revenue loss of MBP-MLE-Limited is to judiciously set the intervals between price changes so that the information is exploited late enough for an accurate estimator, but early enough for the price change to have an



(a) Deviation of $\hat{m} = -50\%$

(b) Deviation of $\hat{m} = 450\%$

Figure II.3: Cumulative revenue of MBP-MLE, MBP-MLE-Limited, and no-learning relative to the upper bound R^* of the optimal pricing-and-learning revenue and 95% confidence intervals.

impact on the total revenue. One possibility is to set the adoptions between price changes to increase exponentially. Theorem II.5 and the ensuing discussion show that such a choice achieves $\mathcal{O}(\ln m_0)$ revenue loss. Note that increasing the length of the exploitation periods over time is similar to other policies proposed in the pricing-and-learning literature (see for example *Broder and Rusmevichientong* 2012; *den Boer* 2015a).

The same logic discussed above is also behind why free learning occurs after the initial price exploration phase in the case of an unknown $x(\cdot)$ function.

II.6 Numerical Study

We compare the performance of MBP-MLE, MBP-MLE-Limited, and a no-learning pricing policy (i.e., MBP policy based on prior parameter values without updating) relative to the optimal policy of the pricing-and-learning problem (II.2.9). We do this by numerically computing the difference between the expected revenues of these heuristics to R^*, which is the optimal expected revenue if the firm knows the parameter vector θ_0.

Figure II.3 shows how the revenues are changing with respect to the initial modeling error (i.e., by how much the initial parameter values are different from the true values). In these examples, we assume that true parameters are $\theta_0 = (p_0, q_0, m_0) = (0.4, 0.6, 100)$. The x-axis represents the percentage deviation $(\hat{\theta} - \theta_0)/\theta_0$ of the initial estimate $\hat{\theta} = (\hat{p}, \hat{q}, \hat{m})$ from the true parameters. We vary the percentage deviation of $\hat{p} + \hat{q}$ from

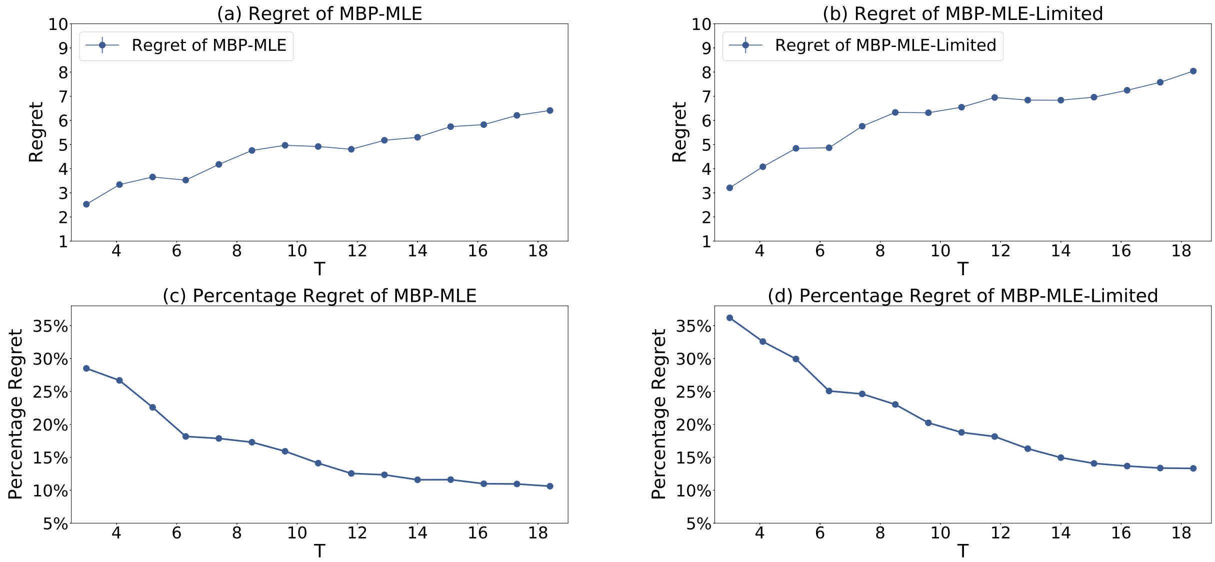


Figure II.4: Cumulative revenue loss and percentage cumulative revenue loss of MBP-MLE and MBP-MLE-Limited.

-90% to 400%, while the deviation of $\hat{m}$ is -50% in (a) and 450% in (b). While the revenue from the optimal policy can be explicitly computed, the revenues from other policies are computed from a simulation with 10^2 trials. Along with the average revenue, we show the 95% confidence intervals. The two panels of Figure II.3 cover all scenarios, namely (overestimate $p_0 + q_0$, overestimate m_0), (overestimate $p_0 + q_0$, underestimate m_0), (underestimate $p_0 + q_0$, overestimate m_0), and (underestimate $p_0 + q_0$, underestimate m_0).

In all the scenarios, the performances of MBP-MLE and MBP-MLE-Limited give us on average 30% more revenue compared with MBP without learning. On the other hand, the performance of a policy that uses only initial estimates degrades sharply as the error gets large. In some instances, such a policy can lose more than 70% of the potential revenue. In contrast, a policy with at most six price changes (MBP-MLE-Limited) based on the data can perform as well as the optimal policy and a policy that requires continuous price changes (MBP-MLE) for most cases except when initial errors are extremely large (around 400%). Even in these cases, MBP-MLE-Limited is significantly better than the no-updating policy. This implies that, if the firm is able to make a few price adjustments after a launch based on the demand data, it can reap substantially more revenue.

Figures II.4 (a) and (b) illustrate how the algorithms perform as T becomes large while keeping m_0 fixed at a finite number. This is a setting that is not considered in our asymptotic regime. We observe that the percentage revenue loss of MBP-MLE and MBP-MLE-Limited, shown in Panels (c) and (d), decrease rapidly. Each dot in the figure is the average revenue loss from a simulation of 10^2 trials. All cases assume that the true parameters are $(p_0, q_0, m_0) = (0.05, 0.1, 160)$ and initial parameters are $(\hat{p}, \hat{q}, \hat{m}) = (0.4, 0.6, 280)$. We also assume that $x(r) = e^{-r}$. From Figure II.4, we clearly see that the revenue loss of policy MBP-MLE-Limited grows faster than that of MBP-MLE.

Note that MBP-MLE does not rely on a prior distribution of the unknown parameters. Therefore, a natural question is: can the performance be improved by a Bayesian estimator that uses a prior distribution? To answer this question, we conduct experiments on a data-driven pricing policy that uses the maximum *a posteriori* (MAP) estimator. The MAP estimator is the parameter value with the highest posterior probability value. Here, the posterior distribution is computed by updating the prior distribution using Bayes' rule after taking into consideration the observed data. We devise a new policy where we use the MAP estimate in the MBP price (in Theorem II.1). Accordingly, we name this new policy MBP-MAP.

In the new experiments, we use MAP to estimate $\beta = (\beta_1, \beta_2, \beta_3)$, and assume a prior distribution for β. Because $\beta_1 > 0$ and $\beta_3 < 0$, we assume the prior of β_1 and $-\beta_3$ follows a gamma distribution with the shape parameter to be $\alpha = 8$. Because the sign of β_2 is free, we assume the prior of β_2 follows a normal distribution $\mathcal{N}(\mu, \mu^2/\alpha)$. In the experiments, we test two cases: (a) the mean of the Bayesian prior is the true value, (b) the mean of the Bayesian prior deviates from the true value by -80%.

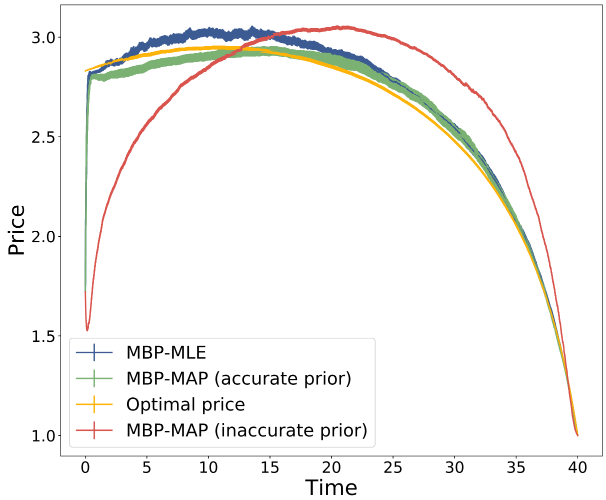


Figure II.5: The true values of the parameters are $m_0 = 150, p_0 = 0.4, q_0 = 0.6, T = 40$. We run 100 experiments of the price sample paths and plot the average with the 95% confidence interval.

Figure II.5 plots the sample average of the price paths (in 100 samples) and 95% confidence intervals under the oracle policy, MBP-MLE, and MBP-MAP. We can see that MBP-MLE (blue curve) and MBP-MAP with an accurate prior (green curve) have average price paths that converge to the optimal price (orange curve). In the short term when there is little data, MBP-MAP with an accurate prior converges faster than MBP-MLE. However, if the prior distribution is inaccurate, MBP-MAP (red curve) has an average price path that is significantly different than the optimal price path. This highlights a weakness of a Bayesian approach in our setting where pricing mistakes can have a lingering effect: that the performance can be very sensitive to the accuracy of the prior when problem size m_0 is relatively small. On the other hand, the quality of prior knowledge has little impact on MBP-MLE.

II.7 Conclusion

This chapter considers how the firm can incorporate learning into pricing decisions for a new product when the demand model parameters are unknown but can be learned from data collected over time. Since firms often do not have sufficient information about adoption behavior and future demand, our work shows that the ability to integrate real-time sales data into the pricing decision can significantly increase revenue.

To develop the mathematical machinery that allows us to capture learning, we propose a new stochastic adoption model, called the Markovian Bass model, that features all the factors affecting state transitions as the generalized Bass model (*Bass*, 1969; *Bass et al.*, 1994). We then show that our Markovian Bass model converges to the Bass model as the market size grows.

Building on this foundation, we study the setting where the seller forms an inference about the unknown Markovian Bass model parameters using maximum likelihood estimation (MLE). We derive the seller's optimal pricing-and-learning policy that strikes a balance between price experimentation and revenue earning. While we are able to characterize the value function of the optimal policy (MBP-MLE-Learning), it is difficult to solve for the optimal policy due to the curse of dimensionality.

Instead, we propose two computationally tractable pricing policies that utilize the ML estimator: MBP-MLE when the retailer has full flexibility to change the price, and MBP-MLE-Limited when the firm must limit the number of its price changes. We show that the worst-case revenue loss of the MBP-MLE grows sub-linearly in the market size. Through a theoretical analysis, we show that the MBP-MLE-Limited achieves the same order of worst-case expected revenue loss as long as the price change intervals are carefully chosen (i.e., the number of adopters between price changes is growing exponentially).

Our framework shows that one can use MLE to derive the optimal learning policy or to develop simple data-driven algorithms with bounded revenue loss when the underlying stochastic process is a continuous time Markov chain. Our result can be applied to other stochastic optimization problems (e.g., pricing, inventory) where the structure and evolution of MLE are well-behaved and leads to state reductions or efficient algorithm development.

CHAPTER III

On the performance of certainty-equivalent pricing

III.1 Introduction

In recent years, many companies have used dynamic pricing as one of the levers to improve their sales revenue. Starting from the 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 the future demand.

In many settings, future demand is uncertain and depends on factors that can change over time. For example, when future demand is driven by a network effect, 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.

These examples of state-varying demand have long been recognized phenomena in marketing and operations management literature. For example, network effects are modeled in the seminal Bass diffusion model (*Bass*, 1969) which is found to fit the empirical demand curve of new products quite well. In this model, sales of a new product are primarily driven by word-of-mouth from satisfied customers. Display effects (demand is high when inventory is high) have been observed in sales data by *Wolfe* (1968); *Smith and Achabal* (1998); *Caro and Gallien* (2012). This effect is attributed to more people noticing the product if there is more inventory. Scarcity effects (demand is high when inventory is low) have been observed experimentally or empirically by *Van Herpen et al.* (2009); *Balachander et al.* (2009); *Cui et al.* (2019); *Cachon et al.* (2018). This effect is attributed to the increase in perceived value because of exclusivity, creating a sense of urgency among customers to “act fast”. With the rise of e-commerce and social media platforms where network or inventory effects could be amplified, it is not surprising that the demand for a product could depend on past sales or inventory.

In order to determine the optimal prices in such settings, the seller must know the distribution of future demand. However, when demand is a complex and state-varying stochastic process, the seller may not have complete demand information. In many settings, the seller’s best available information is the estimate of the average demand in future periods. Indeed, estimating conditional means from data uses standard statistical methodologies (relying on strong results like the law of large numbers), whereas estimating an entire distribution requires a much larger data set and more sophisticated approaches.

In this chapter, we study a periodic-review[1] pricing problem over a finite horizon and with finite inventory when the demand distribution is state-varying. The key features that distinguish our demand model from others in the dynamic pricing literature are that we assume that the future demand and its distribution are state-dependent (where the state variables in our setting are the total past sales and the current inventory level) and that the seller only has limited information about the demand distribution. When demand is state-dependent, a pricing mistake not only reduces the current period revenue, but also changes future demand since the mistake affects cumulative sales and available inventory. Thus, price in one period has a lingering effect on future demand. Furthermore, when there is a limited number of opportunities to change price, the price chosen at each period has persistent implications beyond the current period. Lack of knowledge about the demand distribution makes the pricing decision more difficult and nonoptimal pricing more consequential.

Certainty equivalent (CE) pricing policies are commonly used when the seller has access to the expected demand rather than the entire distribution. Specifically, 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 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 current inventory information at the beginning of each period. Hence, prices are adjusted

[1]Periodic review means that prices can only be changed at the start of each period. While a continuous review of pricing is ubiquitous in analytical models of dynamic pricing, periodic pricing changes are often more appropriate in reality (*Yang and Zhang*, 2014; *Bitran and Mondschein*, 1997). Indeed, periodic pricing schemes are widely observed in practice. For example, many brick-and-mortar stores update their prices weekly as changing prices often requires changing price stickers and are costly to implement.

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 across time are independent and price is reviewed continuously (*Gallego and Van Ryzin*, 1994; *Jasin*, 2014). However, even though the phenomena of state-varying demand and periodic pricing reviews 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. Our work addresses this gap.

A major challenge with this setting is that the state-dependent demand results in non-convex stochastic and deterministic problems that are challenging to analyze. A modeling contribution of our paper is to introduce a general state-dependent demand modeling framework where the certainty equivalent policy is amenable to analysis. Our framework is general in the sense that it includes many of the state-dependent demand models proposed in the literature, such as *Bass* (1969); *Smith and Achabal* (1998); *Shen et al.* (2013); *Smith and Agrawal* (2017).

An initial focus of our analysis is to establish the tractability of solving for the optimal CE policies. The deterministic version of the stochastic problem appears to be difficult to solve at first, due to demand censoring terms in the objective and non-convex constraints. However, through a series of transformations, we show the problem is equivalent to a convex optimization model that has a unique interior solution, and hence can be solved efficiently through interior point methods. Hence, solving for the CE policy is computationally tractable.

Next, we study the analytic performance bounds for the CE policies by comparing their expected revenues against the true (unknown) optimal expected revenue. We do this in two steps. First, we show that for 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. Although a deterministic upper bound can be trivially established in the canonical dynamic pricing setting (e.g., *Gallego and Van Ryzin* 1994), these standard techniques cannot be used in our setting with state-dependent demand and periodic price changes. Hence, we develop a novel induction argument and establish this upper bound using dynamic programming reformulations of the deterministic and stochastic problems.

Second, we show that as the initial inventory and the expected demand are both scaled by m, the gap between the expected revenue of a CE policy (open-loop or closed-loop) and the deterministic upper bound grows in the order $\mathcal{O}(\sqrt{m})$. 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.

In our setting with state-dependent demand, proving the $\mathcal{O}(\sqrt{m})$ upper bound on the revenue loss is challenging since the analysis must apply for any state-dependent distribution (satisfying some regularity conditions). By constructing an appropriate martingale and utilizing the Azuma-Hoeffding inequality, we show 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) proves that re-optimization can reduce the revenue loss from $\mathcal{O}(\sqrt{m})$ to $\mathcal{O}(\log m)$. We show that when demand is state-dependent, this same order reduction cannot be guaranteed for a re-optimization policy. We do this by proving that in our setting the expected revenue loss of both CE policies is lower bounded by $\Omega(\sqrt{m})$. Hence, our $\mathcal{O}(\sqrt{m})$ bound on the expected revenue loss is tight for both CE policies. What this means is that re-optimization has less benefit in our setting compared to its benefit under independent demand. In the latter setting, the reduction to $\mathcal{O}(\log m)$ requires the condition that more inventory strictly improves the revenue (condition $\mu^D > 0$ in Theorem 1 of *Jasin* 2014). However, in our setting of state-dependent demand, more inventory could result in a strictly lower revenue. An example where this could happen is when scarcity boosts sales, so higher inventory results in a lower demand rate.

In our numerical experiments with small m, we show that there is a benefit for re-optimization. This benefit increases as the number of review periods increases or as the conditional mean of future demand becomes more nonlinear in the state variables. Further, both CE policies perform very well, even when m is small (and thus far from the asymptotic limit of $m \to \infty$). Combined with our asymptotic results, this suggests that the CE approach can be used in both large and small markets. To derive a policy that can be used in practice, our policy explicitly takes as an input the number of price changes allowed during a selling season. We show, through large-scale numerical experiments, that just a few price changes are needed 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. A little bit of price flexibility goes a long way. This is reminiscent of *Gallego and Van Ryzin* (1994)'s finding that a fixed price is sufficient to recover the potential benefits of full dynamic pricing at the asymptotic rate $\mathcal{O}(1/\sqrt{m})$ for stationary demand. There is one notable difference. Because future demand depends on price, cumulative sales, and remaining inventory in our model, we show analytically that a fixed-price policy performs poorly.

Finally, 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. We believe this to be the first such result in the literature.

III.1.1 Literature review

In the operations literature, deterministic (CE) 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.* (2013) 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 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). A number of 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 price can be changed continuously. 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. Together, these two conditions allow associated CE problems to be formulated as a linear or convex programs. Thus, theoretical analyses of these settings 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 $\mathcal{O}(\sqrt{m})$[2] 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, a fixed-price CE policy performs close to the optimal policy with continuous price changes. Since then, a number of papers show 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 reoptimization 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 the heuristic. The expected revenue loss of their policy is $o(\sqrt{m})$.

Different from the above settings studied in RM literature, we examine how CE policies perform under general state-dependent demand settings where the seller reviews prices periodically. In particular, we consider the case where demand depends on the cumulative sales and/or on the remaining inventory. Our analysis does not need to assume a specific demand distribution and is general enough to include existing demand settings in the RM literature such as continuous-time Poisson demand arrivals. Accordingly, we contribute to the dynamic literature by providing general conditions under which the CE policy can be an effective alternative to solving the original stochastic optimization problem.

We conclude this section with a table (Table III.1) that positions our work 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

[2]Notation $\mathcal{O}, \Omega, \Theta$ are defined in Section III.1.2.

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

	State-dependent demand	Periodic pricing	Stockout	Partial demand information	Inventory decision
Gallego and Van Ryzin (1994)	No	No	No	No[a]	No
Bitran and Mondschein (1997)	No	Yes	Lost sales	No	Initial
Feng and Gallego (2000)	Yes	No[b]	No	No	No
Shen et al. (2013)	Yes	No	Backlog & lost sales	No	Replenish
Yang and Zhang (2014)	Yes	Yes	Backlog	No	Replenish
This work	Yes	Yes	Lost sales	Yes	Initial

[a] Original paper assumes Poisson demand distribution.
[b] Considers only finitely many price levels.

number of papers with related questions, models, and results.

III.1.2 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 $f(x) = \Theta(g(x))$.

III.2 Modeling framework

We present the *limited information* periodic review pricing problem when the only information about the stochastic per-period demand is its conditional expectation. In this model, 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 periodically at certain price review periods $\{1, 2, \ldots, T\}$. After the firm chooses a price $\pi_t \geq 0$ for period t, a random variable D_t 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. The firm does not know the true distribution of D_t, but it knows the conditional expectation of D_t. Specifically, 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.

The challenge when the firm only knows the conditional expectation of demand is that, if it makes a pricing mistake due to limited information, these mistakes can be costly since future demand is state-dependent. This is because the demand depends on the past sales and the remaining inventory, so any past pricing mistakes can have a lasting effect on future demand. In this paper, we will present pricing policies that only make use of the information on the conditional expectation of demand and analyze their performance in an asymptotic setting (specified in Section III.4.2). In the asymptotic setting, we scale both the expected demand rate and initial inventory by a factor of $m > 0$ while keeping number of price changes T fixed. This means we consider the setting where both demand and inventories are large.

III.2.1 Demand model

We begin by describing the 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.

A distinctive feature of our model is that the per-period demand D_t is a random variable whose distribution may depend on the demand realizations from past periods. However, we assume that conditional on $\mathcal{F}_{t-1}$ and the price π_t, the distribution of D_t only 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 $N_0 = \alpha$.

This feature of the demand model is formalized next.

Assumption III.1. Conditioning on $\mathcal{F}_{t-1}$ and price π_t, the distribution of D_t depends only on π_t, the remaining inventory N_{t-1}, and the cumulative sales $\alpha - N_{t-1}$. Furthermore,

$$\mathbb{E}[D_t \mid \mathcal{F}_{t-1}] = \lambda(N_{t-1}, \alpha) \cdot x(\pi_t) \tag{III.2.1}$$

for some functions λ and x. ◁

The term $\lambda(N_{t-1}, \alpha)$ represents 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

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 III.2: Notation for modeling framework.

the functions $\lambda(\cdot,\cdot)$ and $x(\cdot)$, and that the only information available to the seller about the demand distribution is the functional form of the conditional expectation.

Assumption III.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 III.2 summarizes the notation of our framework, working from (III.2.1) as a primitive.

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

(i) $x : [0,\infty) \mapsto [0,1]$. Moreover, there exists a finite *choke price* π^c where $x(\pi^c) = 0$.

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

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

(iv) $\rho(\pi) \triangleq \pi x(\pi)$ is continuously differentiable in π and $\rho''(\pi)$ exists for all $\pi < \pi^c$.

(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.

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

Assumption III.2(i)-(iv) are standard properties of a price sensitivity function in the revenue management literature. The condition in Assumption III.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. The range $[0,1]$ gives an interpretation of price sensitivity as proportionally "clawing back" on "raw" demand. The existence of the choke price implies that if the price is too high, no one buys. Assumption III.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 known as the virtual value function since it represents the virtual value of the marginal demand resulting from a marginal price change to π. Assumption III.2(iv) implies that the effective revenue rate ρ is a strictly concave function and so has a unique maximizer $\bar{\pi}$ in $[0, \pi^c]$. That is, price $\bar{\pi}$ provides the optimal effective revenue rate. Assumption III.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.

The following examples illustrate how a variety of demand models studied in the literature satisfy the conditions of Assumptions III.1 and III.2.

Example 1 (Sales-dependent demand)**.** The generalized Bass model of *Bass et al.* (1994) and *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

$$\lambda(N_{t-1}, \alpha) = (k - \alpha + N_{t-1}) \left(p + q \cdot \frac{\alpha - N_{t-1}}{k} \right), \tag{III.2.2}$$

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.*, 2013) or linear (*Raman and Chatterjee*, 1995) function. In both these cases, x is consistent with Assumption III.2. Note that λ in (III.2.2) also satisfies Assumption III.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 III.2, but their demand format is in additive form, thus violating Assumption III.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 $\mathbb{E}[D_t \mid \mathcal{F}_{t-1}] = \lambda(N_{t-1})x(\pi)$.[3] The

[3] *Smith and Agrawal* (2017) consider a multi-location inventory model where inventory is sold to

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 modification as follows:

$$\lambda(N_{t-1}) = k(N_{t-1}/(kN_r))^{\beta} \tag{III.2.3a}$$
$$x(\pi) = e^{-\gamma\pi/c_e} - \epsilon_x, \tag{III.2.3b}$$

where k is a market size, N_r and c_e are reference values, and $0 < \beta < 1$, $\gamma > 0$, and $\epsilon_x > 0$. Note that λ is concave, reflecting a diminishing marginal rate of return. Including ϵ_x in (III.2.3b) is a modification of the model in *Smith and Agrawal* (2017) (which assumes $\epsilon_x = 0$) so that a finite choke price exists. Since the choice of ϵ_x is arbitrary, it does not change the results and insights of their paper. We can easily verify that these choices for λ and x satisfy Assumption III.2. $\triangleleft$

We also make the following assumption on the variance.

Assumption III.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 III.3 implies that, relative to the mean, the variance of demand does not become too large. This is not a restrictive assumption. Assumption III.3 is not any stronger than what is assumed in classical dynamic pricing literature where it assumed that demand follows a Poisson or Bernoulli process (*Gallego and Van Ryzin* 1994; *Jasin* 2014) which satisfy Assumption III.3. In fact, we are imposing weaker assumptions than in those works since we are not assuming a specific demand model for our analysis to work. 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 this assumption. If $\sigma = 0$ then demand is deterministic, which is also included in the class of demand models considered this paper.

Example 4. The following are a few distributions that satisfy Assumption III.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,

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).

(b) D_t has a Poisson distribution with mean $\lambda(N_{t-1}, \alpha)x(\pi_t)$. A common demand distribution assumed in dynamic pricing literature is a Poisson process with constant arrival rate λ. This demand model fits our framework since it satisfies Assumptions III.1 to III.3.

(c) D_t is an aggregation of a continuous time Markov chain with transition rate

$$\lambda(N_{t-1}, \alpha)x(\pi_t).$$

III.2.2 The dynamic pricing problem with complete information

We first formulate the dynamic pricing problem when the seller has complete information of the demand process. When the seller does not have explicit information about the demand distribution, the corresponding stochastic optimization model cannot be solved. However, this model later serves as a baseline to evaluate the performance of the certainty-equivalent policies which operate on partial information.

Starting with initial inventory α, the seller chooses a price for each period based on the state. (We call this a periodic-review pricing policy or simply pricing policy.) By Assumption III.1, the conditional distribution of demand D_t given $\mathcal{F}_{t-1}$ depends on the remaining inventory N_{t-1}, and the cumulative sales $\alpha - N_{t-1}$. Therefore, 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

$$V^{\boldsymbol{\pi}}(T) = \mathbb{E}\left[\sum_{t=1}^{T} \boldsymbol{\pi}(N_{t-1}, t)\left(D_t - [D_t - N_{t-1}]^+\right)\right]. \tag{III.2.4}$$

Note that, in our problem setting, the 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 (III.2.4) is taken with respect to a stochastic demand process that is consistent with Assumptions III.1 to III.3.

Remark III.1 (Notational conventions)**.** The superscript $\boldsymbol{\pi}$ of $V^{\boldsymbol{\pi}}(T)$ denotes decisions by the seller. Since we examine how the number of price changes (T) affects the algorithm and resultant profits, we do not suppress T. $\triangleleft$

In light of the properties of the price sensitivity function x, we can recast the seller's decision problem. Assumption III.2(ii) allows us to introduce a new variable $y_t = x(\pi_t)$ called the induced demand intensity for price π_t (or simply *intensity*) at review period t. The price $\pi_t = x^{-1}(y_t)$ 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 $\mathbf{y}$, by Assumption III.1 we have

$$\mathbb{E}\left[D_t \mid \mathcal{F}_{t-1}\right] = \lambda(N_{t-1}, \alpha)\mathbf{y}(N_{t-1}, t), \quad \text{for all } t = 1, \ldots, T. \tag{III.2.5}$$

As the existing literature (e.g. *Gallego and Van Ryzin* 1994) shows, 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

$$V^{\mathbf{y}}(T) \triangleq \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}\left(N_{t-1}, t\right)\right)\left(D_t - [D_t - N_{t-1}]^+\right)\right]. \tag{III.2.6}$$

Discussion of the set of candidate (feasible) intensity policies is needed to complete the description of the seller's problem. 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 policies. The seller's problem is to choose a feasible intensity policy (and thus pricing policy) to maximize the expected revenue, which is equivalent to solving the following problem:

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

We denote the optimal value of this optimization problem ($\mathbf{P}$) by $V^*(T)$. Consistent with our discussion in Remark III.1, parameter T remains an argument of the function $V^*(T)$.

III.3 Certainty-equivalent policies

Note that solving the stochastic pricing problem ($\mathbf{P}$) requires knowing the demand distribution at all states. This is not possible in our setting as the seller only knows the conditional expectation of demand: $\lambda(\cdot, \cdot)$ and $x(\cdot)$. In this section, we introduce pricing policies that only require the limited information in our setting. We refer to these as the certainty-equivalent (CE) policies, since they rely on solving a deterministic counterpart of the stochastic pricing problem ($\mathbf{P}$).

III.3.1 A deterministic optimization model

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

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

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

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

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

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 the deterministic relaxation of ($\mathbf{P}$), where we replace all random variables D_t with their expectations $\lambda(n_{t-1},\alpha)y_t$. Note that ($\mathbf{D}^\dagger$) determines a vector of intensities $y = (y_1, y_2, \dots, y_T)$, whereas ($\mathbf{P}$) finds an intensity policy function $\mathbf{y} : [0,\infty) \times \{1,\dots,T\} \mapsto [0,1]$. Here, $n = (n_1, n_2, \dots, n_T)$ is the vector of remaining inventories under the deterministic demand model. Once a feasible vector y is chosen, the associated feasible vector n is immediately determined.

Note that the objective function ($\mathbf{D}^\dagger$a) in the deterministic model contains censored terms, hence it is non-differentiable. Furthermore, ($\mathbf{D}^\dagger$b) is a non-convex constraint. Hence, at first, the problem appears difficult to solve. However, we will overcome this difficulty by showing that problem ($\mathbf{D}^\dagger$) is equivalent to the following deterministic problem without censoring terms:

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

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

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

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

$$y_t \in [0,1] \text{ for all } t = 1,\dots,T. \tag{$\mathbf{D}$e}$$

We refer to the model above as ($\mathbf{D}$). Note that ($\mathbf{D}$) has an additional constraint ($\mathbf{D}$b) which eliminates solutions (n, y) where the total demand exceeds inventory u. The equivalence between ($\mathbf{D}$) and ($\mathbf{D}^\dagger$) is established in the following theorem:

Theorem III.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**†).

Theorem 1 implies that it suffices to solve problem (**D**) as the deterministic relaxation of the stochastic problem (**P**). This is important since the objective function (**D**a) of the deterministic relaxation (**D**) does not have demand censoring terms (i.e., non-differentiability), which makes it an easier problem to solve. 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 $V^*(T)$.

We will later introduce two CE policies in Section III.3.2, a closed-loop CE policy (CE-CL) and an open-loop CE policy (CE-OL). These two CE policies set the intensity (and equivalently, price) in each period based on solutions to the deterministic model (**D**) for given u and α values. Hence, the complexity of the CE policies depends on the feasibility and computational effort needed to solve the nonlinear optimization problem (**D**). We discuss these properties of (**D**) next.

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 simple transformation:

$$\begin{aligned}
d_1 &= \lambda(u, \alpha) y_1, && \text{(III.3.3a)}\\
d_2 &= \lambda(u - d_1, \alpha) y_2, && \text{(III.3.3b)}\\
d_3 &= \lambda(u - d_1 - d_2, \alpha) y_3, && \text{(III.3.3c)}\\
&\vdots \\
d_T &= \lambda(u - d_1 - d_2 - \ldots - d_{T-1},\ \alpha) y_T. && \text{(III.3.3d)}
\end{aligned}$$

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 ($\mathbf{D'}$):

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

$$\text{s.t.} \quad \sum_{t=1}^{T} d_t \le u \qquad (\mathbf{D'}\text{b})$$

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

The advantage of ($\mathbf{D'}$) is that it is a convex optimization problem with a unique optimal solution, as we establish next. Hence, ($\mathbf{D'}$) can be solved efficiently with commercial off-the-shelf solvers using standard convex optimization algorithm. This means that we can efficiently find the solution of the deterministic counterpart ($\mathbf{D}$), whereas the stochastic problem ($\mathbf{P}$) cannot be solved due to insufficient information about the demand distribution.

Theorem III.2. The following hold:

(i) The objective function ($\mathbf{D'}$a) is jointly concave in d, and the set of all solutions satisfying constraints ($\mathbf{D'}$b)–($\mathbf{D'}$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 ($\mathbf{D'}$) is always feasible since the solution d where $d_t = 0$ for all t is feasible. (Note that by Assumption III.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 Theorem III.3.

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

Theorem III.2 implies that ($\mathbf{D'}$) can be solved efficiently by any standard convex optimization algorithm (e.g., Newton's method) or an off-the-shelf commercial solver. In the next result, we show that the optimal solution to ($\mathbf{D'}$) lies in the interior of the feasible set. This implies that one can deploy interior point methods to determine the optimal solution.

Theorem III.4 (Positive intensity is optimal)**.** If $0 < u \leq \alpha$, then the unique optimal solution d^{D} to ($\mathbf{D'}$) lies in the interior of the feasible set, i.e., $\lambda(u - d_1^{\mathsf{D}} - \ldots - d_{t-1}^{\mathsf{D}}, \alpha) > d_t^{\mathsf{D}} > 0$ for all t.

III.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 ($\mathbf{D}$) which sets the intensity in each period. The fact that the reformulated problem ($\mathbf{D'}$) is well-behaved (Theorem III.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 y^{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 III.1 below describes the CE-OL policy.

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

```
1: procedure OPEN-LOOP CERTAINTY EQUIVALENT PRICING(α, T)
2:     d^D ← optimal solution of (D′) with u = α
3:     for t ← 1 to T do
4:         y_t^D ← d_t^D / λ(α − d_1^D − d_2^D − … − d_{t−1}^D, α)
5:         set intensity y_t^D by offering price x^{−1}(y_t^D)      ▷ set current intensity (price)
6:     end for
7: end procedure
```

On the other hand, 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 as $\mathbf{y}^{\mathsf{CL}}$. 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 $y^{\mathsf{D}} = (y_1^{\mathsf{D}}, y_2^{\mathsf{D}}, \ldots, y_{T-t+1}^{\mathsf{D}})$. 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 III.2 below is a description of the CE-CL intensity policy.

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

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
  end for
end procedure

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 Macbook Pro with an Intel i5 processor, it takes less than 10 seconds to solve ($\mathbf{D}'$) with $T = 22$ using a basic interior point algorithm coded in Python. Note that in ($\mathbf{D}'$), the number of variables is T and the number of constraints is $T + 1$. Neither of the CE policies require solving for the intensity in all possible states. More specifically, the CE-CL policy can be implemented and solved on-the-fly by observing the available inventory at the start of each period (that is, there is no need to anticipate the available inventory).

III.4 Analysis of certainty-equivalent policies

Our goal in this section is to analyze the performance of the two policies proposed in Section III.3.2. The main challenge in analysis is that the demand in period t can depend on the past sales or available inventory. As a result, any pricing mistake in the current period affects current demand and the demand in future periods. Another challenge is that, while our goal is to propose an algorithm that only utilizes partial information (i.e., conditional mean), the performance analysis must apply to all demand distributions satisfying Assumptions III.1 to III.3. The main result of this section 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 $\mathcal{O}(1/\sqrt{m})$.

Our approach in proving the convergence rate is through 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 $V^*(T)$. Recall that $V^{\mathsf{D}}(T)$ 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.

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

The challenge in proving that $V^{\mathsf{D}}(T)$ is an upper bound for $V^*(T)$ in our setting comes from the fact that demands are state-dependent and prices can only be changed periodically.

To see why, consider a situation where the demand rate is a constant λ (independent of the state) and price can be changed continuously. Due to continuous price changes, as soon as the 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. 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 $\mu \geq 0$

$$\begin{aligned} V^\lambda(T) = \mathbb{E}\left(\int_0^T x^{-1}(y_t^\lambda)\mathrm{d}D_t\right) &\leq \mathbb{E}\left(\int_0^T x^{-1}(y_t^\lambda)\mathrm{d}D_t + \mu\left(\alpha - \int_0^T \lambda y_t^\lambda \mathrm{d}t\right)\right) \\ &\leq \max_{y_t : t\in[0,T]} \left(\int_0^T x^{-1}(y_t)\lambda y_t \mathrm{d}t + \mu\left(\alpha - \int_0^T \lambda y_t \mathrm{d}t\right)\right). \end{aligned} \tag{III.4.1}$$

The first inequality is from Langrangian relaxation since we know that the expected demand cannot exceed α. The second inequality is from maximizing pointwise for each t and by Jensen's inequality. Note that the right-hand side of (III.4.1) 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 $\mu \geq 0$.

In our setting with state-dependent demand and periodic price changes, this same approach cannot be used to establish the upper bound. The first issue is that price changes are periodic, so the stochastic objective (III.2.6) has a demand censoring term. This means that the deterministic relaxation ($\mathbf{D}^\dagger$) is a non-convex optimization problem and strong duality does not necessarily hold. Even though we showed the equivalence of ($\mathbf{D}^\dagger$) to the model ($\mathbf{D}$) without censoring (see Theorem III.1), note that constraint ($\mathbf{D}$b) is non-convex so strong duality is still not guaranteed if this constraint is relaxed. A second issue comes from the fact that demand is state-dependent. As a result, the pointwise maximum in (III.4.1) 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:

$$\begin{aligned} R^{\mathsf{D}}(u,T) \quad &\triangleq \quad \max_{y\in[0,1]} \quad x^{-1}(y)\lambda(u,\alpha)y + R^{\mathsf{D}}\left(u - \lambda(u,\alpha)y,\ T-1\right) \\ &\qquad\qquad \text{s.t.} \quad\ \lambda(u,\alpha)y \le u, \end{aligned} \tag{III.4.2}$$

where the base case is $R^{\mathsf{D}}(u, 0) = 0$ for all $u \in [0, \alpha]$. Observe that $R^{\mathsf{D}}(u, T)$ can be thought of as the deterministic revenue-to-go if the remaining inventory is u and there are T periods 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:

$$R^*(u,T) \quad \triangleq \quad \max_{y\in[0,1]} \quad \mathbb{E}_{y,u}\left[x^{-1}(y)\left(D - [D-u]^+\right) + R^*([u-D]^+,\ T-1)\right]. \tag{III.4.3}$$

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, 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 periods 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 when we condition on u, T, the constraint in $\lambda(u, \alpha)y \le u$ in (III.4.2) is linear in y. Using mathematical induction, we can also establish that the objective of (III.4.2) is strictly concave in y due to the concavity assumption on $\lambda(\cdot, \cdot)$. Hence, strong duality holds for the Lagrangian relaxation of (III.4.2). 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 (III.4.3) 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 III.1 (Upper bound)**.** For any $T \geq 1$, $V^*(T) \leq V^{\mathsf{D}}(T)$. More generally, for any $0 \leq u \leq \alpha$, $R^*(u,T) \leq R^{\mathsf{D}}(u,T)$.

We prove this result through mathematical induction on T, starting from establishing the bound for $T = 1$. The complete proof can be found in Appendix B.2.6.

III.4.2 Asymptotic regime

Consider a scaled version of the problem, where we introduce $m \in \mathbb{Z}^+$ as a scaling factor. Thus, we scale the initial inventory to be equal 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 III.4. The conditional expectation of the demand D_t^m has a SIS function λ^m that scales in m such that

$$\lambda^m(N_{t-1}^m, \alpha m) = m\lambda\left(\frac{N_{t-1}^m}{m}, \alpha\right), \tag{III.4.4}$$

where λ is a function that is independent of m and that satisfies Assumption III.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 III.4, together with Assumption III.1, implies that the conditional expectation of demand scales up with m. Note that Assumption III.4 is only required for the proof of asymptotic optimality. Assumption III.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 III.4 holds by simply scaling the demand rate as λm.

In the demand model of Example 1, Assumption III.4 holds if the market size scales as km. Indeed, from (III.2.2), 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 III.4 also holds when the market size scales as km. From (III.2.3a), 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 III.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 $\lambda(N_{t-1}^m, \alpha m) = m\lambda(N_{t-1}^m/m, \alpha)$.

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

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

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

$$V^{\mathbf{y}}(m, T) \triangleq \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}(N_{t-1}^m, t)\right)\left(D_t^m - [D_t^m - N_{t-1}^m]^+\right)\right]. \qquad \text{(III.4.5)}$$

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 $N_t^m = [N_{t-1}^m - D_t^m]^+$, where $N_0^m = \alpha m$ is the scaled initial inventory. For any m, the distribution of D_t^m satisfies Assumptions III.1 to III.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 III.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 III.1 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, due to (III.4.4), $\lambda^m(mn, m\alpha) = m\lambda(n, \alpha)$ 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, $(n^{\mathsf{D},m}, y^{\mathsf{D},m})$ and $(n^{\mathsf{D}}, y^{\mathsf{D}})$, have the property that $n^{\mathsf{D},m} = mn^{\mathsf{D}}$ and $y^{\mathsf{D},m} = y^{\mathsf{D}}$. 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 III.4.3, we show that both expected revenue losses are lower bounded by $\Omega(\sqrt{m})$. Then, in Section III.4.4 we show that both expected revenue losses are upper bounded by $\mathcal{O}(\sqrt{m})$ (i.e., slower than linear). 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 argument 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 demand inventory paths to their deterministic equivalents (see Lemmas III.1 and B.4 below). Assumption III.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.

III.4.3 Lower bound on CE expected revenue loss

It is known that for the open-loop certainty equivalent policy, a lower bound on $V^{\mathsf{D}}(m,T) - V^{\mathsf{OL}}(m,T)$ is $\Omega(\sqrt{m})$ (see Remark 2 in *Jasin* 2014). In the next result, we formally establish that in our setting with state-dependent demand, under the closed-loop certainty-equivalent policy, $V^{\mathsf{D}}(m,T) - V^{\mathsf{CL}}(m,T)$ is also lower bounded by $\Omega(\sqrt{m})$.

Theorem III.5. There exists a distribution satisfying Assumptions III.3 and III.4 such that the expected revenue loss under CE-CL is $V^{\mathsf{D}}(m,T) - V^{\mathsf{CL}}(m,T) = \Omega(\sqrt{m})$.

Jasin (2014) shows that when the demand follows a price-controlled Bernoulli process, then under a regularity condition the re-optimization certainty-equivalent pricing policy can achieve an $\mathcal{O}(\log m)$ bound on the expected revenue loss. The $\mathcal{O}(\log m)$ bounds relies on the independence of demand in each period, which helps with the martingale construction and the tight characterization of dual variables in the certainty equivalent problem. Due to the state-dependence of demand in our setting, we cannot use the arguments of *Jasin* (2014) to show that the CE-CL policy has a revenue loss that grows $\mathcal{O}(\log m)$.

III.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 the order $\mathcal{O}(\sqrt{m})$. Hence our lower bound result implies that, under a setting with state-dependent demand and periodic price reviews, both certainty equivalent policies have an expected revenue loss that is $\Theta(\sqrt{m})$.

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 $u = \alpha$. Hence, even though the conditional expectation of demand is state-dependent in our setting, this lemma implies that the expected demand rate of the open-loop policy converges in expectation to the deterministic optimal demand rate.

Lemma III.1 (Convergence of remaining inventory and SIS)**.** If $n^{\mathsf{D}} = (n_1^{\mathsf{D}}, \ldots, n_T^{\mathsf{D}})$ is the solution to (**D**) when $u = \alpha$, then the following hold:

$$\mathbb{E}\left|\frac{\bar{N}_t^m}{m} - n_t^{\mathsf{D}}\right| = \mathcal{O}\left(1/\sqrt{m}\right), \quad \text{for all } t = 1, \ldots, T, \tag{III.4.6}$$

$$\mathbb{E}\left|\lambda\left(\frac{\bar{N}_t^m}{m}, \alpha\right) - \lambda\left(n_t^{\mathsf{D}}, \alpha\right)\right| = \mathcal{O}\left(1/\sqrt{m}\right), \quad \text{for all } t = 1, \ldots, T. \tag{III.4.7}$$

The proof of this lemma is in Appendix B.2.8. The challenge in the proof lies in the fact that the demands across periods are dependent, so we cannot write the remaining

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

inventory $\bar{N}_t^m$ as the sum of independent random variables and use standard convergence results. We overcome this challenge by constructing an appropriate martingale so that we can apply the Azuma-Hoeffding's inequality for martingales to find the gap between $\bar{N}_t^m$ and its unconditional expectation $\mathbb{E}\left(\bar{N}_t^m\right)$ without knowing the functional form of λ and its unconditional distribution.

With the help from Lemma III.1, 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 (III.4.8) below) is computed, assuming all demands can be sold irrespective of the inventory level. The proof is in Appendix B.2.9.

Lemma III.2 (Convergence of uncensored revenue)**.** The following holds:

$$\left|\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda^m\left(\bar{N}_{t-1}^m, \alpha m\right) y_t^{\mathsf{OL}}\right) - V^{\mathsf{D}}(m,T)\right| = \mathcal{O}\left(\sqrt{m}\right). \tag{III.4.8}$$

Though the bound in Lemma III.2 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 III.1, establishes the asymptotic bound for the expected revenue loss of $\mathbf{y}^{\mathsf{OL}}$ (Theorem III.6 below). Specifically, the proof of the next result (in Appendix B.2.10) shows that the censored revenue $V^{\mathsf{OL}}(m,T)$ converges to the uncensored revenue as m grows large.

Theorem III.6 (Expected revenue loss of open-loop CE policy)**.** The following holds:

$$1 - \frac{V^{\mathsf{OL}}(m,T)}{V^{*}(m,T)} \leq 1 - \frac{V^{\mathsf{OL}}(m,T)}{V^{\mathsf{D}}(m,T)} = \mathcal{O}\left(1/\sqrt{m}\right). \tag{III.4.9}$$

The implication of Theorem III.6 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 III.6 applies for any demand distribution, as long as Assumptions III.1 to III.3 hold.

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 $\mathbf{y}^{\mathsf{CE}}(n,t)$ is Lipschitz continuous in any $n \in [0, \alpha m]$. This is formalized in the following lemma.

[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.

Lemma III.3 (Lipschitz continuous policy)**.** There exists C_y such that, for any $n, n' \geq 0$,

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

This is 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 $\mathbf{y}^{\mathsf{CE}}$ 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 a 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 $\mathbf{y}^{\mathsf{CE}}$ to the deterministic inventory sequence. This is formalized in Lemma B.4, which is stated and proved in Appendix B.2.12. This then allows us to show that the *uncensored* expected revenue under $\mathbf{y}^{\mathsf{CE}}$ has a gap from $V^{\mathsf{D}}(m,T)$ that is $\mathcal{O}(\sqrt{m})$. This is formalized in Lemma B.5, which is stated and proved in Appendix B.2.13. Note that Lemma B.4 and Lemma B.5 are the counterparts of Lemma III.1 and Lemma III.2 for the closed-loop policy.

Hence, as with the open-loop policy, the closed-loop certainty equivalent policy $\mathbf{y}^{\mathsf{CE}}$ is asymptotically optimal to the stochastic periodic pricing problem as the problem scale m grows large. Its proof is in Appendix B.2.14.

Theorem III.7 (Expected revenue loss of closed-loop CE policy)**.** The following holds:

$$1 - \frac{V^{\mathsf{CE}}(m,T)}{V^{*}(m,T)} \leq 1 - \frac{V^{\mathsf{CE}}(m,T)}{V^{\mathsf{D}}(m,T)} = \mathcal{O}\left(1/\sqrt{m}\right). \tag{III.4.10}$$

The asymptotic optimality of the closed-loop policy holds for any demand distribution that satisfies Assumptions III.1 to III.3

III.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 III.1 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. This deterministic upper bound implies asymptotic optimality of the CE policies if the expected demand "path" converges to the deterministic demand path. But in our setting with inventory-dependent demand, this requires first proving that the stochastic inventory levels of the CE policies converge to the optimal deterministic inventory levels. 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. To prove this inventory path convergence, we define a super-martingale with a bounded difference and use a martingale concentration inequality, as seen in the proofs of Lemmas III.1 and B.4. Note that when the demands are independent, it is unnecessary to show the inventory path convergence for the CE policies to be asymptotically optimal.

The second feature is that the prices are reviewed periodically. 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 III.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 once inventory reaches zero. Due to the uncensored demand, the analysis in those continuous price review models is tractable. Perhaps a setting resembling limited price changes 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, they restricted their analysis to policies where the resulting total demand does not exceed inventory almost surely, so there is no demand censoring in the objective. With periodic pricing reviews, reasonable policies could result in lost sales on some demand sample paths. Hence, our analysis of asymptotic optimality needs to hold in the case of demand censoring. We are able to overcome the challenge of demand censoring 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 III.1) and in proving the inventory path convergence of the CE policies (Lemmas III.1 and B.4). 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 III.6 and III.7.

III.5 Extensions

III.5.1 Joint optimization of starting inventory and pricing

We next study 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 c per each unit of inventory. Suppose that the demand distribution is dependent on the starting inventory and is state-dependent, where the state is the current inventory level. Specifically, the demand distribution satisfies Assumptions III.1 to III.3. The seller only knows the conditional expectation of the per-period demand through the functions λ and x.

If the seller knew the distribution of per-period demand, then her goal will be to maximize the expected profit by jointly optimizing the initial inventory and the periodic-review pricing policy. In this case, she will solve a stochastic dynamic optimization problem to decide the initial inventory α and the pricing policy. The expected profit of a decision $(\alpha, \mathbf{y})$ is

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

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, under 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 III.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 III.3 gives a description of the CE policy.

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

```
procedure CERTAINTY EQUIVALENT(T)
    α^CE ← optimal solution of (D′)
    set N_0 = α^CE                                          ▷ set initial inventory
    set prices according to the CE-policy (open-loop or closed-loop) for (α^CE, T)
end procedure
```

Computing the certainty-equivalent policy for a joint inventory and pricing policy is tractable. Recall that in Theorem III.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 (III.4.4). We denote the optimal expected profit as $Q^*(m,T)$ and the expected profit of the certainty-equivalent policy is $Q^{\mathsf{CE}}(m,T)$. As in the case with the certainty-equivalent pricing policies, we show that the expected profit loss under Algorithm III.3 grows sub-linearly in m – this means that our proposed joint decision policy is asymptotically optimal. This is formally established in Theorem III.8. The proof is in Appendix B.3.1.

Theorem III.8 (Expected profit loss of CE policies)**.** The following holds:

$$1 - \frac{Q^{\mathsf{CE}}(m,T)}{Q^*(m,T)} = \mathcal{O}\left(1/\sqrt{m}\right). \tag{III.5.1}$$

This result shows that the CE policy guarantees a close-to-optimal expected profit when the scale of inventory and demand is large. This result is somewhat surprising since $\alpha^{\mathsf{CE}} m$ is not necessarily equal to the optimal initial inventory of the mth stochastic problem (which we denote by $\alpha^* m$). Hence, the fact that the CE policy may choose a different initial inventory implies that the asymptotic optimality in Theorem III.8 does not follow immediately from Theorems III.6 and III.7. But the implication of Theorem III.8 is that when m is large enough, the scaled down initial inventory α^* is close to α^{CE}.

III.5.2 Analysis of a fixed-price policy

When the demand rate is time-stationary and independent, a fixed-price policy (i.e., setting the same price for all time periods) is known to be asymptotically optimal (*Gallego and Van Ryzin*, 1994). We next analyze the performance of such a policy under our problem setting with state-dependent demand.

Given the initial inventory $\alpha \geq 0$, we first define the fixed-price policy $\mathbf{y}^{\mathsf{SP}}$. If α is sufficiently large, the fixed-price policy fixes a price corresponding to intensity $\bar{y}$, where $\bar{y} \in [0,1]$ is the unique maximizer of the revenue function, i.e., $\bar{y} \triangleq \arg\max_{y\in[0,1]} x^{-1}(y)y$. In other words, if the inventory constraint is nonbinding, the policy chooses the intensity that maximizes the current period revenue only, without considering the effects of inventory and sales on demand. 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"):

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

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 $t \in \mathcal{T}$. Note that $\bar{y}^{\mathrm{so}}$ can be found by fixed point iteration.

Mathematically, given any initial inventory $\alpha \geq 0$, the fixed-price policy $\mathbf{y}^{\mathsf{SP}}$ is defined for every $(n,t) \in (0,\alpha] \times \mathcal{T}$ as:

$$\mathbf{y}^{\mathsf{SP}}(n,t) = y^{\mathsf{SP}} \triangleq \begin{cases} \bar{y}, & \text{if } \alpha \geq \sum_{t=1}^{T} \lambda\left(n_{t-1}^{\bar{y}}, \alpha\right)\bar{y}, \\ \bar{y}^{\mathrm{so}}, & \text{otherwise.} \end{cases} \tag{III.5.2}$$

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

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

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

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

$\mathbf{y}^{\mathsf{SP}}$. Specifically,

$$V^{\mathsf{D}',\alpha}(T) \triangleq \sum_{t=1}^{T} x^{-1}\left(y^{\mathsf{SP}}\right)\lambda(n_{t-1}^{\mathsf{SP}},\alpha)y^{\mathsf{SP}}, \tag{III.5.3}$$

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

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

procedure FIXED POLICY(T)
  $\alpha^{\mathsf{SP}} \leftarrow$ optimal solution of (**S**)
  set $N_0 = \alpha^{\mathsf{SP}}$ ▷ set initial inventory
  set prices with FIXED PRICING(α^{SP}, T)
end procedure

procedure FIXED PRICING(α, T)
  $y^{\mathsf{SP}} \leftarrow \bar{y}$ or $\bar{y}^{\mathsf{so}}$ based on cases in (III.5.2) for α
  for $t \leftarrow 1$ **to** T **do**
    set intensity y^{SP} by offering price $x^{-1}(y^{\mathsf{SP}})$ ▷ set current intensity (price)
  end for
end procedure

We next state the main result of this subsection which describes the performance of the fixed-price policy under our setting. 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 $(m\alpha^{\mathsf{SP}}, \mathbf{y}^{\mathsf{SP}})$ as $Q^{\mathsf{SP}}(m,T)$. For any $\alpha \ge 0$, we denote $V^{\mathsf{SP},\alpha}(m,T)$ as the expected revenue under the stochastic model of the fixed-price policy $\mathbf{y}^{\mathsf{SP}}$ with initial inventory $m\alpha$, and $V^{*,\alpha}(m,T)$ as the expected revenue under the optimal pricing policy with initial inventory $m\alpha$.

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

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

then $V^{*,\alpha}(m,T) - V^{\mathsf{SP},\alpha}(m,T) = \Omega(m)$. If (i)–(ii) hold for $\alpha = \alpha^{\mathsf{SP}}$, then $Q^{*}(m,T) - Q^{\mathsf{SP}}(m,T) = \Omega(m)$.

The proof is in Appendix B.3.2. Condition (i) of Proposition III.2 implies the myopic optimal intensity $\bar{y}$ is not the optimal first-period price for deterministic model $V^{\mathsf{D}}(T)$. Condition (ii) implies that initial inventory is sufficiently large. Proposition III.2 shows that both profit loss and revenue loss of a fixed-price policy grows at least linearly in the scaling factor m.

One may think that the reason that the fixed-price policy performs poorly is that it may not start with the optimal initial inventory, i.e., $\alpha^{\mathsf{SP}} \neq \alpha^*$. However, Proposition III.2 shows that, regardless of the initial inventory level, the profit loss grows at least at a linear rate in the scaling factor as long as the initial inventory level is sufficiently large. This shows that the inability to adjust the price results in a much greater loss when demand depends on inventory and cumulative sales.

In contrast, the certainty-equivalent policies allow the seller to adjust price, even if the price sequence is static (i.e., open-loop policy). Thus, whether the future demand is driven by past sales or by inventory availability or by both, the seller can account for the current revenue as well as the future revenue when setting prices. The difference in the fixed pricing policy and a certainty-equivalent policy can be demonstrated when $T = 2$. Proposition III.2 shows that a fixed price results in a loss growing at least at a linear rate in the scaling factor. By contrast, the revenue and profit loss of certainty-equivalent policies are order $\mathcal{O}(\sqrt{m})$, which implies asymptotic optimality (see Theorems III.6 to III.8). This means, even a single opportunity to change the price (based on the sales and inventory) can substantially reduce the revenue or profit loss.

III.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). We first illustrate the analytic properties of the deterministic value function $V^{\mathsf{D}}(T)$. In Section III.6.2, we show the CE-OL and CE-CL converge fast numerically and can achieve close-to-optimal performance even in instances with a small scaling factor m. In Section III.6.4, we experiment on the number of price changes and demonstrate the value of increased flexibility in pricing.

III.6.1 The deterministic revenue $V^{\mathsf{D},\alpha}(T)$ and the initial inventory problem

We illustrate the deterministic revenue function $V^{\mathsf{D},\alpha}(T)$ with a concrete example. Following Example 3, we choose price sensitivity function $x(\pi) = e^{-\gamma\pi} - c_x$. We consider a case where the demand is influenced by both the 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,

$$\lambda(n, \alpha) = \left(w\lambda^{(1)}(n, \alpha) + (1 - w)\lambda^{(2)}(n, \alpha)\right) \Delta t, \tag{III.6.1}$$

where $\lambda^{(1)}(n, \alpha) = ((n - \alpha^2 + 1)/N_r)^\beta$ (cf. (III.2.3a)) and

$$\lambda^{(2)}(n, \alpha) = (1 - (\alpha - n)) (p + q(\alpha - n))$$

(cf.(III.2.2)), 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 (III.2.3a) 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 (III.6.1) is jointly concave in (n, α). The parameters used in this example are $(p, q, N_r, \beta, \gamma, c_x, T, \Delta t) = (0.4, 0.6, 25, 0.6, 0.001, 0.01, 10, 2)$.

Figure III.1 plots the optimal value function $V^{\mathsf{D},\alpha}(T)$ as a function of the initial inventory α with different weights w of the SIS function (III.6.1). Without loss of generality, we normalize the demand so that we impose the constraint $\alpha \in [0, 1]$. The figure illustrates that $V^{\mathsf{D},\alpha}(T)$ is concave in α, which agrees with item (ii). When $w = 0$, only the network effect (the positive effect of sales on demand) comes into play, and so it is optimal to serve the market fully ($\alpha = 1$). When $w = 1$, only scarcity effects are felt and the optimal initial inventory is $\alpha = 0.68$. When $w = 0.5$ (network effect, saturation effect, and scarcity effect are all present), the optimal choice of inventory is $\alpha = 0.84$. This complex example with all three effects present shows that we should choose $\alpha < 1$ in the presence of a scarcity effect of inventory.

III.6.2 Revenue loss of the certainty-equivalent policy

We next illustrate the performance of the CE policies on the demand pattern considered in Section III.6.1. We set $w = 0.5$ in (III.6.1) 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 $\mathbf{y}^{\mathsf{CE}}$ is based on re-optimizing (**D**) in each period with updated inventory levels. The policy $\mathbf{y}^{\mathsf{OL}}$ does not re-optimize the revenue in each period but sets time-varying prices.

We vary the inventory and demand scaling factor m from 100 to 3000, with discretizations shown in the horizontal axis of Figure III.2. For each m, we randomly generate

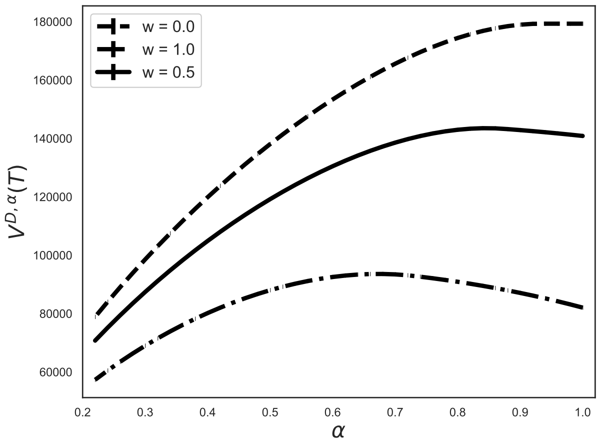


Figure III.1: The deterministic revenue function (**D**) plotted against the initial inventory α, for different values of w in (III.6.1).

2×10^4 demand sample paths following a bounded support Poisson distribution; we implement the dynamic pricing policies $\mathbf{y}^{\mathsf{OL}}$ and $\mathbf{y}^{\mathsf{CE}}$, 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 $\bar{V}^{\mathsf{CE}}(m,T)$, are the approximations for the expected revenue of the certainty-equivalent policies, $V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{OL}}}(m,T)$ and $V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{CE}}}(m,T)$ 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 an 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 $V^{\mathsf{CE}}(m,T)$ 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 $(V^{\mathsf{D}}(m,T) - \bar{V}^{\mathsf{CE}}(m,T))/V^{\mathsf{D}}(m,T)$, which are shown as the points in Figure III.2. The figure also shows the 95% confidence intervals of the revenue loss bound. From Theorem III.7, 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 III.2. We further observe that the revenue losses by implementing both $\mathbf{y}^{\mathsf{OL}}$ and $\mathbf{y}^{\mathsf{CE}}$ are very small ($\sim 0.15\%$ when $m = 3000$). This implies that, for a product with 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).

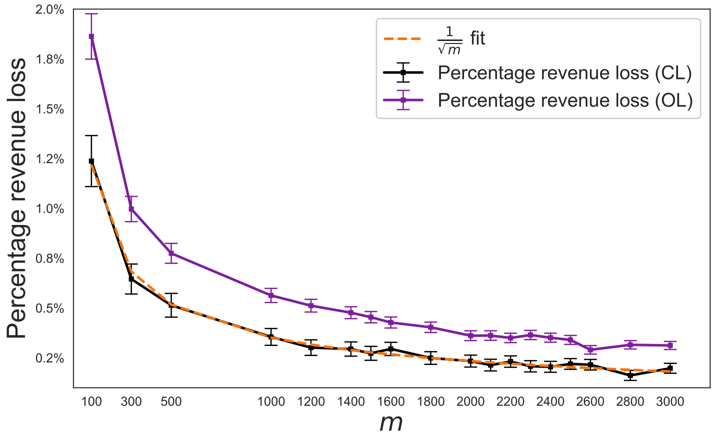


Figure III.2: Upper bound on the percentage revenue loss of the certainty-equivalent policies against the optimal value of the stochastic problem. The fixed-price policy has a bound on percentage revenue loss that is at least 30% (not shown in graph).

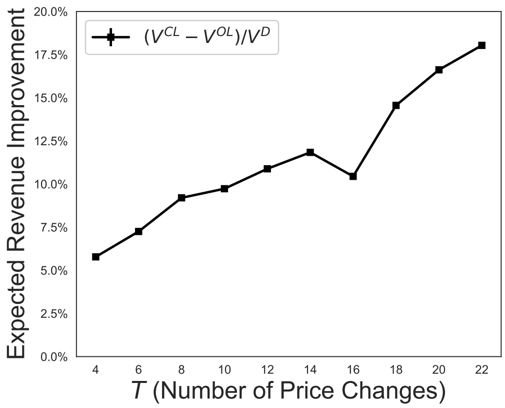


Figure III.3: Value of resolving by increasing number of price changes

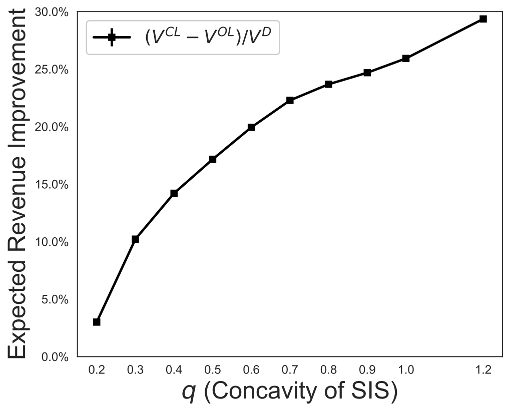


Figure III.4: Value of resolving by increasing the concavity of the SIS function

III.6.3 The benefit of re-optimization in non-asymptotic setting

In contrast to the open-loop policy CE-OL, the closed-loop policy CE-OL requires solving the deterministic problem (**D**) with the updated information in each period where the firm can change price. One important question is how beneficial is reoptimization. From Theorem III.5, we know that in an asymptotic regime, CE-CL does not improve the revenue loss of CE-OL (which grows at the rate of $\sqrt{m}$) in any meaningful way. To see where the benefit of reoptimization exists in non-asymptotic settings, we conduct a numerical study comparing the two policies.

Figure III.3 shows how the gain from re-optimization over the open-loop policy changes as the number of price changes T increases. In this example, demand follows a Poisson distribution and the Bass SIS function defined in (III.2.2) with p fixed at 0.01, when $q = 1.0$, and $k = 20$. The figure shows that more frequent re-optimization is beneficial as more opportunities of adjust prices reduce the probability of an early stock-out during the selling horizon and generate more revenue out of the remaining inventory. We note that benefit of re-optimization has an increasing trend if there are more opportunities for changing prices

Figure III.4, on the other hand, shows how the gain from re-optimization 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 function is highly non-linear and concave, the static $\mathsf{CE\text{-}OL}$ current price typically deviates more from the optimal policy. For instance, if the SIS function follows a Bass function, as defined in (III.2.2), the second-order derivative with respect to inventory decreases with q, where q is the imitation parameter in Bass's terminology. This means that as q increases (more people imitate), the seller will lose significant revenue by not re-optimizing (**D**).

To understand the performance gap between $\mathsf{CE\text{-}OL}$ and $\mathsf{CE\text{-}CL}$ policy intuitively, consider the following. The closed-loop policy re-optimizes the price in each period. As a result, given state information, under the closed-loop policy the expected demand does not exceed the remaining inventory. This implies that the conditional expected inventory follows $\mathbb{E}(N_t \mid \mathcal{F}_{t-1}) = N_{t-1} - \lambda(N_{t-1}, \alpha)\mathbf{y}^{\mathsf{CL}}(N_{t-1}, t)$, which moves similarly to the deterministic path $n_t^{\mathsf{D}} = n_{t-1}^{\mathsf{D}} - \lambda(n_{t-1}^{\mathsf{D}}, \alpha)y_t^{\mathsf{D}}$. This means that, as m increases, $\hat{N}_t^m/m$ will be quite close to n_t^{D}. In contrast, $\mathsf{CE\text{-}OL}$ policy does not guarantee that the conditional expectation of N_t/m 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 $\mathbb{E}(N_t \mid \mathcal{F}_{t-1}) \neq N_{t-1} - \lambda(N_{t-1}, \alpha)y_t^{\mathsf{OL}}$. Although, as m increases, the normalized demand will not overshoot its expectation more than $1/\sqrt{m}$, we can intuitively see the absolute deviation of the normalized inventory $\bar{N}_t^m/m$ from n_t^{D} under the open-loop policy might be larger than that of the closed-loop policy because of the former policy's lost sales in expectation. Furthermore, as the number of review periods T increases or the concavity of λ increases, the deviation of $\bar{N}_t^m/m$ from n_t^{D} will be worse because demand variance increases in T and the difference in expected demand increase in $-\frac{\partial^2 \lambda}{\partial n^2}$. This, in turn, increases the probability that there will be lost sales under the open-loop policy. However, we note that y_t^{OL} and y_t^{CL} are close as m increases.

III.6.4 Revenue loss due to limited price changes

The certainty-equivalent policies we consider are discrete-time policies that assume that the underlying demand is modeled as a discrete-time process. Hence, an interesting question to ask is: how much revenue can the discrete-time 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 that is 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 (III.6.1). That is, conditional on current inventory level n, the probability of having one sale during a time period of length $o(t)$ is

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

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 $(\alpha^{\mathsf{CE}}, \mathbf{y}^{\mathsf{CE}})$ 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 III.3 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 III.6.2, 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 is almost as good as continuous price changes.

These results provide numerical evidence that a few price changes is good enough to capture the revenue from changing price continuously (which is very costly in practice). A small number of prices go a long way. 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 III.3, we found that the deterministic optimal policy is Lipschitz continuous in the remaining inventory. This means the difference in the two policies is

Table III.3: 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	4	5	10	17	22	35	45
95 % CI lower bound	70.3%	95.6%	97.9%	98.4%	99.4%	99.6%	99.8%	99.9%	100.0%
Expected normalized revenue	70.3%	95.6%	97.9%	98.5%	99.4%	99.7%	99.8%	100.0%	100.0%
95 % CI upper bound	70.3%	95.7%	98.0%	98.5%	99.4%	99.7%	99.9%	100.0%	100.0%

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.

III.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. 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 cumulative 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 forms of scarcity (in case of luxury or fad items) or billboard effect and can be found in many markets today. Moreover, we consider a periodic review pricing policy, which is commonly practiced in reality. Our work also provides a justification for the current literature studying dynamic pricing for a product with non-stationary demand using a deterministic approach.

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 regret rate of $\Theta(\sqrt{m})$, when compared with the optimal policy. However, in a non-asymptotic setting, CE-CL performs consistently better than CE-OL, especially when the number of price change periods increases and when the conditional expectation of demand is highly nonlinear in past cumulative sales and available inventory. This result highlights the importance of re-optimization in the face of sales and inventory dependent demand. We then extend our results to the case where the seller chooses initial inventory along with 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 of dynamic pricing under sales and inventory dependent demand against 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 optimality even in cases where the scaling factor is not large. Furthermore, we show that significant revenue improvement can be achieved by just a few price changes.

There are several future directions for our work. One is to extend the framework to the multi-product case where those products share the same market. Another extension is to consider strategic customers. The customers can strategically wait until there is a discount. *Sapra et al.* (2010) touch 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. Another direction is to incorporate learning into our model. In this chapter, we assume that the conditional expectation of the demand is known. It is possible to approximate the expectation using available data throughout the selling horizon.

CHAPTER IV

Valuing influence

IV.1 Introduction

Influencer marketing — promoting products or services through social-media *influencers* — has been a popular practice in recent years. Influencers are individuals who share their impressions of a particular product category (such as fashion, technology, gaming, travel) in order to shape the opinions of their followers. They create content on social media platforms (e.g., YouTube, Instagram, and Facebook) with active followers who consume this content. Influencers earn money through their reputation, which includes being consistent and true to their values. Companies leverage these well-earned reputations in exchange for money, to target customers who follow influencers. Companies recognize the value of influencers in spreading valuable information about their products, particularly information that they may not be able to credibly communicate themselves. This marketing practice allows companies to attract interest in a product from pre-selected active customers. According to a report from Wondershare (*Brown*, 2020), a platform providing video-making software, the top ten richest content creators on YouTube earned between $10 million to $15.5 million in 2019. InfluencerDB, a platform that collects data on influencer marketing, reports that spending on Instagram influencers alone exceeded $5 billion in 2018 (*Vardhman*, 2019).

Although companies spend large sums on influencer marketing, 61% of them find it difficult to identify effective influencers for their product campaign *MediaKix* (2019). An effective influencer increases brand awareness and converts awareness into sales. Influencers are typically paid by companies by a flat rate per "post", where a post is a video, picture, or message on a social media platform. Revenue-sharing arrangements exist but are much rarer in practice.[1] However, the *value* of a post is hard to assess. The following

[1] Verified through personal communication with social media influencers.

questions are still largely unanswered:

(i) how lucrative is it for a company to work with a given influencer?

(ii) who is the ideal influencer to post about a given product?

(iii) what attributes of an influencer make her more or less valuable to a company?

This work aims to develop an analytical model designed to answer these and related questions.

In order to value influence, several factors must be considered. First, and most obviously, is the number of followers that form the audience of the influencer. In industry practice, influencers are categorized based on the number of followers they have: nano (500–5,000 followers), micro (5,000–30,000 followers), power (30,000–500,000 followers), and celebrity (more than 500,000 followers). According to a survey of more than 2,500 influencers from January 2019 to March 2019 released by *Klear* (2019), the pay rate per post for a celebrity is more than 20 times that of a nano-influencer.

The size of the following is a commonly used factore for valuing influencers in practice. Their impact is somewhat well-understood and therefore not a major focus of our research. Our interest is in the informational value of influencers; that is, the nature of how followers update their beliefs upon receiving the influencer's opinion. To our knowledge, the information value of influencers has not been carefully explored in the literature. Below, we describe two dimensions of an influencer's ability to change the beliefs of her followers: reputation and charisma.

Followers shape their beliefs of the product based on the influencer's historical *reputation*; that is, how revealing her posts are regarding the true nature of the product. An influencer who always endorses products, even if they turn out to have bad quality, is considered to have a bad reputation. The post of such an influencer reveals little information. The positive review of an influencer who is more selective in their praise carries more information.

Moreover, different followers may put different weights on the influencer's opinion versus the opinions of others. *Devotees* only listen to the influencer, while *skeptics* learn from both the influencer and other followers. The influencer's *charisma* is the fraction of her followers who are devotees. Because of the existence of skeptics, followers cannot be considered to be independent decision makers. The revenue generated from the influencer is affected by the dynamic belief processes of followers, which evolve from the influencer's post and the actions of other followers. The greater the charisma of an influencer, the less dependency there is among the decisions of her followers.

In this on-going project, we want to provide a framework for valuing an influencer based on her reputation and charisma. To our knowledge, this is a far more comprehensive model of influence than commonly used in practice, yet our results remain tractable and intuitive. Accordingly, our results provide practical guidance for companies when selecting influencers. Additionally, our modeling framework allows us to identify the features of a company's "ideal" influencer.

A key factor that makes our model insightful – and challenging to analyze – is the complexity of the interactions between followers who dynamically form their beliefs. We handle this complexity through a diffusion approximation, which makes the comparison among different influencers possible. Moreover, for the optimal signaling scheme that maximizes the expected revenue of the company, we show that it suffices to consider the strategy where the influencer restricts to two possible signals. This is achieved by formulating the problem into a semi-infinite linear program. This implies that using a complicated strategy that selects a signal among infinitely many possible options is not helpful in improving revenue.

IV.2 Connections to existing literature

Our model and results are closely related to (i) the marketing and economics literature on strategic communication and information transmission from firms to customers, and (ii) operations and computer science literature on algorithmic information design and optimal signaling mechanisms.

The problem of how firms use information to manipulate the opinions of customers has been extensively studied in the marketing literature. Traditionally, word-of-mouth is endogenously created by customers who experience the product. In contrast, a growing body of literature investigates how a firm's exogenous manipulation of word-of-mouth can be lucrative. *Godes and Mayzlin* (2009) provide empirical evidence that firms can exogenously create word-of-mouth to drive sales of a product with low awareness. This practice is more effective when using less loyal customers or acquaintances of target customers to spread the word because their opinions are considered to be less biased and more credible. It is the first work that empirically shows that word-of-mouth can be partially controlled by firms. Our works continues in this vein, as we examine the phenomena of how companies can use opinion leaders (influencers) with independent credibility and messaging power to impact the evolving beliefs of followers in a social context.

Different from our setting, many marketing papers have explored alternative mechanics for shaping the opinions of their customers. A prime example is the manipulation of online reviews by firms. For example, *Mayzlin et al.* (2014) use review data on TripAdvisor and

Expedia to identify firm manipulation, adding positive reviews to their hotels and negative reviews to their competitors. *Mayzlin* (2006) and *Dellarocas* (2006) propose theoretical models to study the persuasiveness of online reviews, knowing that some of the reviews are manipulated by firms. *Chakraborty and Harbaugh* (2014) develop a theoretical model to show that puffery, which is the claim about product strengths made by the company itself without providing evidence, conveyed to a buyer can be useful in increasing revenue. The intuition behind their result is that when the initial purchase probability is low, puffery increases customers' purchasing probabilities because it emphasizes some attributes of the product that are of interest to certain group of customers.

However, today's consumers are becoming even more savvy. They are aware that anonymous information and online opinions can be manipulated. This partially explains the rise in prominence of social network influencers whose job is to have credibility and authenticity in the online space. The goal of the firm is not to manipulate the message of influencers, but precisely the opposite, to select influencers whose genuine and unmanipulated messaging lines up with the interests of the firm.

Influencers, however, do not operate in a vacuum. It is also important to consider the effect that their followers have on each other. *Kozinets et al.* (2010) studies word-of-mouth marketing among bloggers and finds that, after the initial campaign by the blogger (analogous to an influencer), so-called network co-production of the marketing message by the blog readers can shape the narrative in a social network. Using empirical evidence, *Hamami* (2019) finds that product reviews of early adopters can change the beliefs about product quality in later-arriving customers. Effects of this type are captured in our model by the existence of skeptics who "shape" the message of the influencer and must be considered in valuing an influencer's post.

Our framework also contributes to the economics literature on information transmission and persuasion. The economics literature has long understood the importance of credibility in communication. For instance, *Crawford and Sobel* (1982) propose the classic cheap-talk game where there is an informed signal sender and an uninformed receiver who will react to the sender's signal. In a cheap talk game, there always exists a "babbling" equilibrium that no one has any incentive to communicate anything meaningful and informative. Thereafter, the generalizations of the cheap-talk game includes multiple senders (*Gilligan and Krehbiel*, 1989; *Krishna*, 2001; *Battaglini*, 2002; *Ambrus and Lu*, 2010), multiple equilibria and equilibrium selection (*Farrell*, 1993; *Chen*, 2011; *Kartik*, 2009; *Chen et al.*, 2008; *Gordon*, 2011), and so on (see *Sobel* (2013) for a comprehensive survey). Many traditional methods for influencing word-of-mouth marketing may fall into the "cheap talk" paradigm. These methods have diminishing impact as the notion of online credibility has become more visible. Thus, the cheap talk setting does not as readily apply to our problem because (i) both the influencer (signal sender) and the company need not be better informed than their customers of whether the product will have a bigger or smaller market (before demand is realized), and (ii) the value of an influencer who endorses products online stems from her credibility (she loses the long-run reputation for being a cheap talk sender).

Accordingly, models of credible communication, as studied in the Bayesian persuasion literature, take on greater relevance for studying word-of-mouth marketing. In a Bayesian persuasion model, the sender commits to a signaling mechanism that stochastically determines her messaging. The uninformed receiver is a rational Bayesian learner who decides on an action after learning from the sender's signal. The signaling mechanism can be interpreted as how much information the sender wants to share with the uninformed receiver. *Kamenica and Gentzkow* (2011) and *Rayo and Segal* (2010) provide conditions under which the persuasion is effective, i.e., the sender achieves higher expected payoff by committing to a signaling strategy. Our work is closely related to the persuasion literature because the influencer's reputation cannot be changed in the short run, an implementation of the notion of commitment. The company benefits from sponsoring an influencer only if the reputation of the influencer can "persuade" more followers to make purchases.

Since the seminal work of *Kamenica and Gentzkow* (2011), the study of Bayesian persuasion has developed rapidly. The most relevant strand to our setting are generalizations to multiple receivers. This literature can be divided into two settings, whether the sender sends private signals or public signals. *Arieli and Babichenko* (2019) and *Candogan and Drakopoulos* (2020) study the optimal signaling mechanism in the setting that the sender sends private (possibly different) signals to different receivers. Our focus is on public signaling mechanisms. In practice, a social media influencer sends public messages to all of her followers. This public display of "content" is the key driver of follower engagement.

For Bayesian persuasion literature with public signals, *Alonso and Câmara* (2016a) and *Alonso and Câmara* (2016b) consider the public signaling mechanism in a voting setting. The most significant distinction of our model is that we allow certain group of receivers not only to learn from the signal sent by the sender, but also purchasing behavior of others in the social network. *Candogan* (2019) studies a persuasion game where receivers are socially connected via a network. He proposes tractable algorithms to solve for the static equilibrium. Each receiver's utility depends on the state of the world and the decisions of others in the network. Our work differs from his by making the receivers' utilities and belief processes micro-founded. Specifically, we model decision dependency among receivers via their dynamic learning behavior. We often see in practice that purchases within a group of friends is more like a contagion than a simultaneous-move game amenable to equilibrium analysis. To the best of our knowledge, we are the first to consider a Bayesian persuasion sender (the influencer) followed by sequentially arriving receivers (followers) that can learn from one another.

The sequential arrival of receivers shares commonalities with several papers in a growing literature in operations management that employs Bayesian persuasion as a modeling framework. *Lingenbrink and Iyer* (2019) design optimal signaling mechanisms in a queuing setting. The sender sends private signals about the queue length to sequentially arriving receivers. They formulate the problem as an infinite linear program and analyze the optimal solution in the queue's steady-state distribution. *Lingenbrink and Iyer* (2018) use a two-period model to illustrate the efficacy of public signals in online retail. They propose the optimal signaling strategy of the firm as the solution to a fractional knapsack problem.

A common criticism of Bayesian persuasion models is to put doubt on the sender's ability to credibly commit to her signaling mechanism and, moreover, questions whether any given signaling strategy can be implemented in practice. This criticism particularly applies to the operations management literature, who, by nature, focus on practical aspect of implementation. Our model does not make the (arguably) strong assumption that firms can commit to and implement an arbitrary signalling strategy. Indeed, influencer marketing precisely allows for the firm to leverage alternative signaling mechanisms via influencers. The seller in our setting does not have the ability to credibly provide their desired signal to its target customers. The influencer is the intermediary to carry the message with her long-term reputation. We take the seller's perspective to analyze whether it is profitable to choose a particular influencer with certain features (i.e., reputation, charisma, number of followers, and effective period of the signal).

CHAPTER V

Conclusion and future work

Decision making in uncertain, history-dependent, and non-stationary markets is the theme of my doctoral research. In the information age, we see more and more communication between companies and their customers, and among customers themselves. It is common to see markets with these three features. My dissertation provides answers and practical solutions to Rvenue Management (RM)/Supply Chain Management (SCM) problems when the company faces those complications (uncertainty, history dependency, and non-stationarity) in the market condition. It also inspires me to explore a broader area: how would the complications change the decisions in a supply chain with risk dependency or information sharing in the supply chain network? Nowadays, companies in the supply chain also have stronger linkages through many technologies (e.g., blockchain). This is the next step in my research agenda.

Overall, the goal of my research is to extend our understanding of the “socia” component in RM/SCM problems: relationships among customers and firms, as well as customers themselves. In those relationships, people (firms and customers) choose what to reveal to one another and how to learn from one another. I call this a “relationship-driven” approach to Operations Management (OM) research. The core of relationship-driven OM is as follows: one decides what to share with each other and how to react to (by operations decisions or information strategies) what is shared. These relationships have become prominent because of shifting technologies and the sheer amount and complexity of information available in recent years. Nowadays, the Covid-19 pandemic has further changed the nature of human interactions. This brings research opportunities to relationship-driven OM: how will the new relationship among people impact operations decisions? Does a lack of human connection caused by social distancing lead less information sharing? What is the result? Exploring more into relationship-driven OM problems is my current and future research focus. I believe it to be a powerful lens into the important questions facing our discipline. Moreover, I am always excited to see the linkage between the real world and my research and potentially to contribute to the real practice of RM/SCM in this new age.

APPENDICES

APPENDIX A

Proofs of chapter II

Table of Contents

CHAPTER

A.1 Algorithms

Algorithm A.1 Numerically solve HJB equation (II.3.3) for V

Require: Step size dt, horizon length $T = Ndt$, model parameters (p_0, q_0, m_0), $x(\cdot)$, termination criteria ϵ, max iterations N_{iter}

Ensure: Function $V : \{0, 1, ..., m_0\} \times [0, T] \to \mathbb{R}^+$

$V(m_0, :) \leftarrow 0$, $V(:, 0) \leftarrow 0$ ▷ *Set boundary conditions*
for all $t \in \{dt, 2dt, \ldots, Ndt\}$ **do**
  for all $d \in \{m_0 - 1, m_0 - 2, \ldots, 1\}$ **do**
    $k \leftarrow 0$, $\nu_0 \leftarrow V(d, t - dt)$ ▷ *ν is the estimate for $V(d, t)$*
    $\nu_{-1} \leftarrow \nu_0 + 2\epsilon$
    while $|\nu_k - \nu_{k-1}| > \epsilon$ **and** $k \leq N_{\mathsf{iter}}$ **do** ▷ *Find ν using fixed point iteration*
      $r \leftarrow \inf\left\{r : r \geq -\frac{x(r)}{x'(r)} - V(d+1, t) + \nu_k\right\}$ ▷ *Solve (II.3.2) for r^**
      $\nu_{k+1} \leftarrow V(d, t - dt) + dt(m_0 - d)(p + q\frac{d}{m_0})\frac{x(r)^2}{x'(r)}$ ▷ *Numerically solve (II.3.3) for ν*
      $k \leftarrow k + 1$
    end while
    $V(d, t) \leftarrow \nu_k$
  end for
end for

Algorithm A.2 Numerically solve HJB equation (A.11) for V^{MLE}

Require: Step sizes dt, dp, dq, horizon length $T = Ndt$, model parameters m_0, $x(\cdot)$, termination criteria ϵ, max iterations N_{iter}, a large integer M

Ensure: Function $V^{\mathsf{MLE}} : \{0, 1, ..., m_0\} \times [0, T] \times [0, \infty) \times [0, \infty) \to \mathbb{R}^+$

$V^{\mathsf{MLE}}(m_0, :, :, :) \leftarrow 0$, $V^{\mathsf{MLE}}(:, 0, :, :) \leftarrow 0$ ▷ *Set boundary conditions*

for all $t \in \{dt, 2dt, \ldots, Ndt\}$ **do**
for all $d \in \{m_0, m_0 - 1, \ldots, 0\}$ **do**
$k \leftarrow 0$, $\nu_0(:,:) \leftarrow V^{\mathsf{MLE}}(d, t - dt, :, :)$ ▷ *$\nu(:,:)$ is the estimate for $V^{MLE}(d, t, :, :)$*
$\nu_{-1} \leftarrow \nu_0 + 2\epsilon$
while $\|\nu_k - \nu_{k-1}\| > \epsilon$ **and** $k \leq N_{iter}$ **do** ▷ *Find $\nu(:,:)$ using fixed point iteration*
for all $p \in \{0, dp, \ldots, Mdp\}$ **do**
for all $q \in \{0, dq, \ldots, Mdq\}$ **do**
Compute η_p, η_q from (A.8), (A.9)
$\nabla_d V \leftarrow V^{\mathsf{MLE}}(d + 1, t, p + \eta_p, q + \eta_q) - \nu_k(p, q)$
if $p \in [dp, (M - 1)dp]$ and $q \in [dq, (M - 1)dq]$ **then**
Use centered difference to estimate $\nabla_p V$ and $\nabla_q V$ from ν_k, i.e.,
$\nabla_p V \leftarrow (\nu_k(p + dp, q) - \nu_k(p - dp, q)) / (2dp)$
$\nabla_q V \leftarrow (\nu_k(p, q + dq) - \nu_k(p, q - dq)) / (2dq)$
else
Use forward or backward difference to estimate $\Delta_p V$ or $\Delta_q V$ from ν_k
end if
$r \leftarrow \inf \left\{ r : r \geq -\frac{x(r)}{x'(r)} - \nabla_d V + \nabla_p V \eta_p + \nabla_q V \eta_q \right\}$ ▷ *Solve* (A.10) *for r^**
$\nu_{k+1}(p, q) \leftarrow V^{\mathsf{MLE}}(d, t - dt, p, q) + dt(m_0 - d)(p + q\frac{d}{m_0})\frac{x(r)^2}{x'(r)}$ ▷ *Numerically solve* (A.11) *for ν*
end for
end for
$k \leftarrow k + 1$
end while
$V(d, t, :, :) \leftarrow \nu_k(:, :)$
end for
end for

A.2 Proofs

A.2.1 Proof of Proposition II.1

Proof. We will use the arguments adapted from Chapter 11 ("Density dependent population processes") in the book *Ethier and Kurtz* (2005) to prove (II.2.7). We follow the proof idea used in the book, but the results we cite are well established lemmas/theorems in the literature. We will use the prefix EK to denote the sections and results in the *Ethier and Kurtz* (2005) book.

In the proof below, we decompose the difference $\frac{D_t^{r,m_0}}{m_0} - F_t^r$ into a martingale divided by m_0 and a term that diminishes as m_0 grows. The martingale term converges to zero almost surely by Doob's martingale convergence theorem. In order to show the variance of $\frac{D_t^{r,m_0}}{m_0}$ decreases in the order of $1/m_0$ as m_0 increases, we use a continuous time diffusion process (contains Brownian motion) to approximate the asymptotic distribution of $\frac{D_t^{r,m_0}}{m_0}$. Then we directly compute the asymptotic variance using Itô's isometry. The following is the detailed proof.

First, we introduce some notations. Let Z_λ be an exponentially distributed r.v. with mean $1/\lambda$. Let $Y := \{Y(t), t \geq 0\}$ be a standard Poisson process with intensity 1. Let Y_j be the jth inter-arrival time of Y. Note that Y_j has the same distribution as Z_1. We also define $\tilde{Y} := \{Y(t) - t, t \geq 0\}$, which is a "centered" Poisson process with mean zero. Let Z_j^{r,m_0}, $1 \leq j \leq m_0$, be the jth inter-adoption time in $D^{r,m_0} = \{D_t^{r,m_0}, t \geq 0\}$ and t_{j-1} be the time that cumulative adoption hits $j-1$.

Define the function $A(y) := (1-y)(p_0 + q_0 y)$, $y \in [0,1]$. Note that $\xi(j) = m_0 A(j/m_0)$ is the portion of the adoption rate unaffected by the price when the state of process D^{r,m_0} is j. We use the fact that we can write Z_λ in terms of Z_1 as $Z_\lambda = \frac{1}{\lambda} Z_1$. Hence, we can write $\{D_t^{r,m_0}, t \geq 0\}$ in terms of Y by letting

$$Z_j^{r,m_0} \triangleq \sup \left\{ t \geq 0 : \int_0^t m_0 A\left(\frac{j-1}{m_0}\right) x(r_{s+t_{j-1}}) \mathrm{d}s \leq Y_j \right\}.$$

Then, we know $\{D_t^{r,m_0}, t \geq 0\}$ can be constructed via the standard Poisson process Y with

$$D_t^{r,m_0} = Y\left(\int_0^t m_0 A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s) \mathrm{d}s \right).$$

The above transformation is commonly seen in the literature to construct martingales and it is also shown in Theorem 4.1 of Chapter 6 in *Ethier and Kurtz* (2005).

Therefore, using the newly defined processes and function, we have

$$\frac{D_t^{r,m_0}}{m_0} = \frac{1}{m_0} Y\left(\int_0^t m_0 A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s\right) - \int_0^t A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s + \int_0^t A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s$$
$$= \frac{1}{m_0}\tilde{Y}\left(m_0 \int_0^t A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(\pi_s)\mathrm{d}s\right) + \int_0^t A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s. \tag{A.1}$$

First, since $p_0 > 0$ and $q_0 > 0$ (from our model assumption), we have that the quadratic function $A(y)$ is bounded above by $\bar{A} := p_0 + \frac{(q_0-p_0)^2}{4q_0}$ for any $y \in [0,1]$. Therefore,

$$\left|\frac{1}{m_0}\tilde{Y}\left(m_0 \int_0^t A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s\right)\right|$$
$$\leq \left|\frac{1}{m_0}\sup_{0\leq u\leq t}\tilde{Y}\left(m_0 \int_0^u A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s\right)\right| \leq \left|\frac{1}{m_0}\sup_{u\leq t}\tilde{Y}\left(m_0\bar{A} \int_0^u x(r_s)\mathrm{d}s\right)\right|, \tag{A.2}$$

Sending m_0 to infinity on both sides of (A.2), we have

$$\lim_{m_0\to\infty}\left|\frac{1}{m_0}\tilde{Y}\left(m_0 \int_0^t A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s\right)\right| \leq \lim_{m_0\to\infty}\left|\frac{1}{m_0}\sup_{u\leq t}\tilde{Y}\left(m_0\bar{A} \int_0^u x(r_s)\mathrm{d}s\right)\right| = 0. \tag{A.3}$$

The right hand side of (A.3) is zero almost surely by Doob's martingale convergence theorem.

Therefore, by (A.1) and the definition of F_t^r in (II.2.1), we have

$$\left|\frac{D_t^{r,m_0}}{m_0} - F_t^r\right| = \left|\frac{1}{m_0}\tilde{Y}\left(m_0 \int_0^t A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s\right) + \int_0^t \left[A\left(\frac{D_s^{r,m_0}}{m_0}\right) - A\left(F_s^r\right)\right] x(r_s)\mathrm{d}s\right|$$
$$\leq \underbrace{\left|\frac{1}{m_0}\tilde{Y}\left(m_0 \int_0^t A\left(\frac{D_s^{r,m_0}}{m_0}\right) x(r_s)\mathrm{d}s\right)\right|}_{\Delta_1} + \underbrace{\int_0^t \left|A\left(\frac{D_s^{r,m_0}}{m_0}\right) - A\left(F_s^r\right)\right| x(r_s)\mathrm{d}s}_{\Delta_2}. \tag{A.4}$$

To bound Δ_2, note that $A'(y) = q_0 - p_0 - 2yq_0$. Hence, we have

$$\Delta_2 \le \max_{y\in[0,1]} |A'(y)| \times \int_0^t \left| \frac{D_s^{r,m_0}}{m_0} - F_s^r \right| x(r_s)\mathrm{d}s \le |q_0 + p_0| \int_0^t \left| \frac{D_s^{r,m_0}}{m_0} - F_s^r \right| x(r_s)\mathrm{d}s. \quad \text{(A.5)}$$

Thus, substituting (A.5) into (A.4) and let $X(t) := \int_0^t x(r_s)\mathrm{d}s$, we have

$$\left| \frac{D_t^{r,m_0}}{m_0} - F_t^r \right| \le \Delta_1 + |q_0 + p_0| \int_0^t \left| \frac{D_s^{r,m_0}}{m_0} - F_s^r \right| \mathrm{d}X(s).$$

By Gronwall's inequality, we have

$$\left| \frac{D_t^{r,m_0}}{m_0} - F_t^r \right| \le \min\left\{ \Delta_1 e^{|q_0+p_0| \int_0^t x(r_s)\mathrm{d}s}, 2 \right\}. \quad \text{(A.6)}$$

According to (A.3), we have $\Delta_1 \to 0$ almost surely as $m_0 \to \infty$. Taking m_0 to infinity on both sides of (A.6), we have

$$\lim_{m_0\to\infty} \left| \frac{D_t^{r,m_0}}{m_0} - F_t^r \right| = 0 \quad \text{almost surely,}$$

proving the first part of the proposition.

We next analyze the convergence of the variance of D^{r,m_0}. To do this, we define the new stochastic process $V_t^{r,m_0} := \sqrt{m_0}\left(\frac{D_t^{r,m_0}}{m_0} - F_t^r\right)$. We also define $\{V_t^r, t \ge 0\}$ to be stochastic process satisfying the following stochastic differential equation:

$$\mathrm{d}V_t^r = A'(F_t^r)x(r_t)V_t^r\mathrm{d}t + \sqrt{A(F_t^r)x(r_t)}\mathrm{d}W_t \quad \text{(A.7)}$$

where $\{W_t, t \ge 0\}$ is the standard Brownian motion (i.e., mean is zero, variance is t). Note that the solution of (A.7) is

$$V_t^r = \int_0^t e^{\int_s^t A'(F_u^r)x(r_u)\mathrm{d}u} \sqrt{A(F_s^r)x(r_s)}\mathrm{d}W_s. \quad \text{(A.8)}$$

We can directly use EK Theorem 2.3 in Chapter 11 (p.458) (or e.g., *Kurtz* (1971) Theorem 3.1, equation (1.5) in *Norman et al.* (1974)) to get the following result: For a given t, we have that V_t^{r,m_0} converges to V_t^r in distribution as $m_0 \to \infty$. Therefore, we can find the asymptotic variance of V_t^{r,m_0} by Itô's isometry, and it is equal to

$$\begin{aligned}
&\mathrm{Var}\left(V_t^r\right) \qquad\qquad (A.9)\\
&= \int_0^t \left(e^{\int_s^t A'(F_u^r)x(r_u)\mathrm{d}u}\sqrt{A(F_s^r)x(r_s)}\right)^2 \mathrm{d}s\\
&= F_t^r(1-F_t^r) + (1-F_t^r)\\
&\cdot \frac{2q_0/p_0\left[(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s - 1 + e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}\right] + (q_0/p_0)^2\left(1-e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}\right)^2}{\left(1+q_0/p_0 e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}\right)^3 + e^{(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}}\\
&\le F_t^r(1-F_t^r) + (1-F_t^r)\frac{2q_0/p_0\left[(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s - 1 + e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}\right] + (q_0/p_0)^2}{e^{(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}}\\
&\le F_t^\pi(1-F_t^r) + (1-F_t^r)\alpha\left(\frac{t}{e^t}\right) \qquad\qquad (A.10)
\end{aligned}$$

for some $\alpha > 0$ independent of m_0. By the definition of V_t^{r,m_0}, we have $\mathrm{Var}\left(V_t^{r,m_0}\right) = m_0 \mathrm{Var}\left(\frac{D_t^{r,m_0}}{m_0}\right)$. Therefore, for any $t \ge 0$, we conclude that the asymptotic variance of $\frac{D_t^{r,m_0}}{m_0}$ decreases with rate $1/m_0$.

□

A.2.2 Proof of Lemma II.2

Proof. We will drop the subscript θ_0 from $\mathbb{E}_{\theta_0}$ for simplicity of notation. Note that we have

$$\mathbb{E}\left|\frac{D_t^{r,m_0}}{m_0} - F_t^r\right| \le \underbrace{\mathbb{E}\left|\frac{D_t^{r,m_0}}{m_0} - \mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right)\right|}_{(a)} + \underbrace{\left|\mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right) - F_t^r\right|}_{(b)}.$$

To prove the lemma, we will prove that (a) has an upper bound that is $\mathcal{O}\left(1/\sqrt{m_0}\right)$, while (b) has an upper bound that is $\mathcal{O}\left(1/m_0\right)$.

We first bound (a). Fixing time t, we consider the adoption states at time t of each individual within the population of size m_0. We denote their adoption states as $\zeta_i(t)$ for $i = 1, 2, ..., m_0$. If $\zeta_i(t) = 1$, then individual i has adopted the product by time t, and $\zeta_i(t) = 0$ otherwise. Hence, $D_t^{r,m_0} = \sum_{i=1}^{m_0} \zeta_i(t)$, where D_t^{r,m_0} is the number of adoptions by time t.

Since the population is homogeneous, then $\zeta_1(t), \zeta_2(t), \ldots, \zeta_{m_0}(t)$ are *a priori* identically distributed. We next derive an expression for their mean. Let us define $F_t^{r,m_0} := \mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right)$. Since $D_t^{r,m_0} = \sum_{i=1}^{m_0} \zeta_i(t)$, we know that $\frac{1}{m_0}\sum_{i=1}^{m_0} \mathbb{E}\left(\zeta_i(t)\right) = \mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right) = F_t^{r,m_0}$. Since the population is homogeneous, this means that $\mathbb{E}\left(\zeta_i(t)\right) = \Pr\left(\zeta_i(t) = 1\right) =$

F_t^{r,m_0} for all $i = 1, \ldots, m_0$.

Let $\mathcal{X} := \{\zeta_1(t), \ldots, \zeta_{m_0}(t)\}$ be the set of adoption states, which are identical Bernoulli random variables with mean F_t^{r,m_0}. Note that $\frac{1}{m_0} D_t^{r,m_0} = \frac{1}{m_0} \sum_{i=1}^{m_0} \zeta_i(t)$ is the sample average of a random sample (with size m_0) taken without replacement from $\mathcal{X}$. Hoeffding inequality can be used to bound the deviation of the sample average from its mean when sampling is done without replacement (*Bardenet et al.*, 2015). Therefore, we can use Hoeffding inequality to bound (a). Specifically, for any $\epsilon > 0$,

$$\mathbb{P}\left\{\left|\frac{D_t^{r,m_0}}{m_0} - \mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right)\right| > \epsilon\right\} = \mathbb{P}\left\{\left|\frac{1}{m_0}\sum_{i=1}^{m_0} \zeta_i(t) - F_t^{r,m_0}\right| > \epsilon\right\} \leq 2\exp\left(-2m_0\epsilon^2\right).$$

Hence, we have

$$\begin{aligned}
\mathbb{E}\left|\frac{D_t^{r,m_0}}{m_0} - \mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right)\right| &= \int_0^\infty \mathbb{P}\left\{\left|\frac{D_t^{r,m_0}}{m_0} - \mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right)\right| > \epsilon\right\} \mathrm{d}\epsilon \\
&\leq \int_0^\infty 2e^{-2m_0\epsilon^2} \mathrm{d}\epsilon = \frac{2}{\sqrt{2m_0}}\sqrt{\pi} = \mathcal{O}\left(\frac{1}{\sqrt{m_0}}\right),
\end{aligned}$$

thus, proving the bound for (a).

We next bound (b). Let us define $f_t^{r,m_0} := \frac{d}{dt} F_t^{r,m_0} = \frac{1}{m_0}\frac{d}{dt}\mathbb{E}\left(D_t^{r,m_0}\right)$. Hence, recalling that $\lambda(\cdot,\cdot)$ defined in (II.2.3) is the adoption rate function, we have

$$\begin{aligned}
f_t^{r,m_0} &= \frac{1}{m_0}\mathbb{E}\left[\lambda(D_t^{r,m_0}, r_t)\right] && \text{(A.11)} \\
&= \mathbb{E}\left[\left(1 - \frac{D_t^{r,m_0}}{m_0}\right)\left(p_0 + q_0\frac{D_t^{r,m_0}}{m_0}\right) x(r_t)\right] \\
&= \left(p_0\mathbb{E}\left(1 - \frac{D_t^{r,m_0}}{m_0}\right) + q_0\mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0} - \left(\frac{D_t^{r,m_0}}{m_0}\right)^2\right)\right) x(r_t) \\
&= \left(p_0\left(1 - F_t^{r,m_0}\right) + q_0\left(F_t^{r,m_0} - \mathbb{E}\left[\left(\frac{D_t^{r,m_0}}{m_0}\right)^2\right]\right)\right) x(r_t) \\
&= \left(p_0\left(1 - F_t^{r,m_0}\right) + q_0\left(F_t^{r,m_0} - \left(F_t^{r,m_0}\right)^2 - \mathrm{Var}\left(\frac{D_t^{r,m_0}}{m_0}\right)\right)\right) x(r_t). && \text{(A.12)}
\end{aligned}$$

Dividing both sides of (A.12) by $(1-F_t^{r,m_0})(p_0+q_0F_t^{r,m_0})$, we have

$$\begin{aligned}\frac{f_t^{r,m_0}}{(1-F_t^{r,m_0})(p_0+q_0F_t^{r,m_0})} &= x(r_t)\left[1-\frac{q_0\mathrm{Var}\left(\frac{D_t^{r,m_0}}{m_0}\right)}{(1-F_t^{r,m_0})(p_0+q_0F_t^{r,m_0})}\right] \\ &= x(r_t)\left[1-\frac{q_0}{m_0}\frac{F_t^{r,m_0}+\mathcal{O}(1)\left(\frac{t}{e^t}\right)}{p_0+q_0F_t^{r,m_0}}\right],\end{aligned} \tag{A.13}$$

where the last equality follows from (A.10).

The differential equation (A.13) is similar to the deterministic Bass model (II.2.1), except with a modified market effort function. Hence, modifying (II.2.2) results in

$$\mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right)=F_t^{r,m_0}=\frac{1-\exp\left(-(p_0+q_0)\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s\right)}{1+\frac{q_0}{p_0}\exp\left(-(p_0+q_0)\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s\right)}. \tag{A.14}$$

Then, from (II.2.2) and (A.14),

$$\begin{aligned}&\left|\mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right)-F_t^r\right| \\ &=\left|\frac{1-e^{-(p_0+q_0)\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s}}{1+\frac{q_0}{p_0}e^{-(p_0+q_0)\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m}}\right)x(r_s)\mathrm{d}s}}-\frac{1-e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}}{1+\frac{q_0}{p_0}e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}}\right| \\ &=\left|\int_{\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s}^{\int_0^t x(r_s)\mathrm{d}s}\frac{(1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)X}}{\left(1+q_0/p_0e^{-(p_0+q_0)X}\right)^2}\mathrm{d}X\right| \\ &\le\frac{(1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)\bar{x}^l t}}{(1)^2}\left|\int_{\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s}^{\int_0^t x(r_s)\mathrm{d}s}\mathrm{d}X\right|,\end{aligned}$$

where the last inequality follows from Assumption II.1 (ii).

Hence,

$$\left|\mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right) - F_t^r\right| \le \frac{(1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)\bar{x}^l t}}{(1)^2}\left|\int_0^t \frac{q_0}{m_0}\frac{F_s^{r,m_0} + \mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}} x(r_s)\mathrm{d}s\right|$$
$$\le (1+q_0/p_0)(p_0+q_0)\frac{\bar{x}^u t}{e^{(p_0+q_0)\bar{x}^l t}}\frac{1}{m_0} = \frac{t}{e^t}\mathcal{O}\left(\frac{1}{m_0}\right),$$

where the last inequality follows from $\frac{q_0F_s^{r,m_0}+q_0\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}} = \mathcal{O}(1)$ and Assumption II.1 (ii), proving that (b) has an upper bound that is $\mathcal{O}(1/m_0)$.

□

A.2.3 Proof of Theorem II.1

Proof. We can write the value function $V(d,t)$ by enumerating the outcomes after δt time units. Hence, for any $d \in \{0,1,\ldots,m_0-1\}$ and $t \in (0,T]$, we have

$$V(d,t) = \max_{r\in(-\infty,\infty)}\left\{(r+V(d+1,t-\delta t))\,\lambda(d,r)\delta t + V(d,t-\delta t)\,(1-\lambda(d,r)\delta t) + o(\delta t)\right\},$$

where $\lambda(\cdot,\cdot)$ is defined in (II.2.3). On both sides of the equation, we subtract $V(d,t-\delta t)$, divide by δt, then take the limit as δt approaches zero. This results in the HJB equation

$$\frac{\partial V}{\partial t} = \max_{r\in(-\infty,\infty)}\left\{r\lambda(d,r) + [V(d+1,t) - V(d,t)]\,\lambda(d,r)\right\}. \tag{A.15}$$

We will show existence of a unique solution $V(\cdot,\cdot)$ to (A.15) at the end of this proof.

We next derive the optimal solution $r^*(d,t)$ to the right-hand side of (A.15). Denote the objective function of the right-hand side of (A.15) by $J(r,d,t)$. Since $\lambda(d,r) = \xi(d)x(r)$, then

$$\frac{\partial J}{\partial r} = \lambda(d,r)x'(r)\left(r + \frac{x(r)}{x'(r)} + V(d+1,t) - V(d,t)\right). \tag{A.16}$$

Note that $x'(r) < 0$ (Assumption II.1(iii)) and since $d \le m_0 - 1$, we have $\lambda(d,r) > 0$. Therefore, the first order condition $\frac{\partial J}{\partial r} = 0$ is satisfied by $r = r^*(d,t)$, where $r^*(d,t)$ is defined in the theorem as the solution to (II.3.2). Note that (II.3.2) has a unique solution. This is because, rearranging (II.3.2) as

$$0 = -r - \frac{x(r)}{x'(r)} - V(d+1,t) + V(d,t), \tag{A.17}$$

the right-hand side is strictly decreasing in r (Assumption II.1(iv)), implying that there is a unique root to the equation (A.17).

We next show that $r^*(d,t)$ is the unique maximizer of $J(r,d,t)$ for any (d,t). Using the fact that $V(d+1,t) - V(d,t) = -r^* - \frac{x(r^*)}{x'(r^*)}$, we have

$$\left.\frac{\partial^2 J}{\partial r^2}\right|_{r=r^*} = \frac{\xi(d)}{x'(r)}\left[2x'(r)^2 - x(r)x''(r)\right]. \tag{A.18}$$

Since $2x'(r)^2 - x(r)x''(r) \geq 0$ (Assumption II.1(iv)), and $x'(r) < 0$ (Assumption II.1(iii)), it follows that $\left.\frac{\partial^2 J}{\partial r^2}\right|_{r=r^*} \leq 0$. Hence $r^*(d,t)$ is the unique maximizer of the right-hand side of the HJB equation (A.15) and is therefore the unique optimal price given state (d,t). Finally we can use the equation $\frac{\partial J}{\partial r} = 0$ where $r = r^*(d,t)$, to reformulate (A.15) as (II.3.3).

To complete the proof, we will show that there exists a unique solution V to the HJB equation (A.15). By Theorem VII.T3 (page 208) in *Brémaud* (1981), a unique solution exists if we can replace $\max\limits_{r\in(-\infty,\infty)}$ in (A.15) with $\max\limits_{r\in U_t}$ where U_t is a compact set, and if $r\lambda(d,r)$ and $\lambda(d,r)$ are continuous and uniformly bounded in r and d. To show the first condition, note that (II.3.2) implies that

$$|r^*(d,t)| \leq \left|\frac{x(r^*(d,t))}{x'(r^*(d,t))}\right| + |V(d+1,t) - V(d,t)|.$$

Note that $\lambda^c(r) = m_0(p_0+q_0)x(r)$ is an upper bound for the adoption rate in our system at any state (d,t) and price r. Therefore, $|V(d+1,t) - V(d,t)|$ can be loosely bounded by the optimal T-period expected revenue in a system where the adoption rate is $\lambda^c(r)$. By *Gallego and Van Ryzin* (1994), this expected revenue has a deterministic upper bound $J^D = Tm_0(p_0+q_0)\sup\limits_{r\in(-\infty,\infty)} rx(r)$. By Assumption 1(v), $rx(r)$ is bounded by a finite C_x, so J^D is finite. Furthermore, Assumption 1(iv) implies that $|x(r)/x'(r)|$ is bounded for any r. Hence,

$$u := \sup_{r\in(-\infty,\infty)}\left|\frac{x(r)}{x'(r)}\right| + J^D.$$

is a finite value that bounds the magnitude of $r^*(d,t)$. Hence, we can replace the maximization in (A.15) with $\max\limits_{r\in U_t}$ where $U_t = [-u,u]$. This fulfills the first condition.

To satisfy the remaining conditions, we need to show that $r\lambda(d,r)$ and $\lambda(d,r)$ in (A.15) are continuous and uniformly bounded in r,d.

First, note that $\lambda(d,r) = x(r)\xi(d)$, where $\xi(d) = (m_0-d)(p_0+\frac{d}{m_0}q_0)$. By a change of variables $y = d/m_0$, we can write $\xi(y) = m_0(1-y)(p_0+q_0y)$ which attains a maximum value of $\frac{m_0}{4q_0}(p_0+q_0)^2$ when $y = \frac{1}{2} - \frac{p_0}{2q_0}$. Therefore, $\xi(d) \leq \frac{m_0}{4q_0}(p_0+q_0)^2$ for any d. Furthermore, from Assumption 1(ii), we have $\lambda(d,r) \leq \bar{x}^u \frac{m_0}{4q_0}(p_0+q_0)^2$.

To check whether $r\lambda(d,r) = rx(r)\xi(d)$ is uniformly bounded, note that from Assumption 1(iv), there exists a unique maximizer of $rx(r)$. We define it as $r^\#$. Therefore, we know that $r\lambda(d,r)$ is continuous and uniformly bounded by $r^\# x(r^\#)\frac{m_0}{4q_0}(p_0+q_0)^2$ for any r, d.

Thus, there exists a unique solution to the HJB equation (A.15).

□

A.2.4 Lemma A.1 and proof

The following lemma provides monotonicity properties of the value function V with respect to the state variables (d,t) and the demand model parameters $\theta_0 = (p_0, q_0, m_0)$.

Lemma A.1. The value function $V(d,t;\theta_0)$ has the following properties:

(i) $V(d,t;\theta_0)$ is monotone increasing in $t \in [0,T]$,

(ii) $V(d,t;\theta_0)$ is monotone decreasing in d for $d > m_0\left(\frac{1}{2} - \frac{p_0}{2q_0}\right)$

(iii) $V(d,t;\theta_0)$ is monotone increasing in p_0, q_0 and m_0 for any $(d,t) \in \{0,1,\ldots,m_0\} \times [0,T]$.

Proof. We prove the three parts of the lemma below.

1. From (II.3.3), $\frac{\partial}{\partial t}V(d,t)$ is nonnegative due to the assumption that $x'(r) < 0$ for all r (Assumption II.1 (iii)). Therefore, for all $d \in \{0,1,2,...,m_0\}$, $V(d,t)$ is monotone increasing in $t \in [0,T]$.
2. To prove the monotonicity of V with respect to d, we temporarily treat d as a continuous variable. Then by (A.15), since $\frac{\partial V(d,t)}{\partial t} = J(r,d,t)\big|_{r=r^*(d,t)}$, we have that

$$\frac{\partial^2 V(d,t)}{\partial d \partial t} = \frac{\partial J(r,d,t)}{\partial r}\frac{\partial r}{\partial d}\Big|_{r=r^*(d,t)} + \frac{\partial J(r,d,t)}{\partial d}\Big|_{r=r^*(d,t)} \tag{A.19}$$

Note that the first term in the RHS is zero, hence

$$\begin{aligned}\frac{\partial^2 V(d,t)}{\partial d \partial t} &= \frac{\partial \lambda(d,r)}{\partial d}\left[r + V(d+1,t) - V(d,t)\right]\Big|_{r=r^*(d,t)} \\ &+ \lambda(d,r)\frac{\partial[V(d+1,t)-V(d,t)]}{\partial d}\Big|_{r=r^*(d,t)} \\ &= -\frac{x(r^*(d,t))^2}{x'(r^*(d,t))}\left(q_0 - p_0 - 2q_0\frac{d}{m_0}\right)\end{aligned}$$

$$+\frac{\partial[V(d+1,t)-V(d,t)]}{\partial d}(m_0-d)\left(p_0+q_0\frac{d}{m_0}\right)x(r^*(d,t)),$$

where the second equality follows from (II.3.2) and from the definition of $\lambda(\cdot,\cdot)$. If we define $g(d,t) := \frac{\partial V(d,t)}{\partial d}$, then $g(d,0)=0$ and

$$\frac{\partial g}{\partial t}+\left[(m_0-d)\left(p_0+q_0\frac{d}{m_0}\right)x(r^*(d,t))\right]g$$
$$=-\frac{x(r^*(d,t))^2}{x'(r^*(d,t))}\left(q_0-p_0-2q_0\frac{d}{m_0}\right)+\frac{\partial V(d+1,t)}{\partial d}(m_0-d)\left(p_0+q_0\frac{d}{m_0}\right)x(r^*(d,t)), \tag{A.20}$$

which is a linear differential equation. Solving this differential equation using standard techniques, results in

$$\frac{\partial V(d,t)}{\partial d}\underbrace{e^{(m_0-d)(p_0+\frac{d}{m_0}q_0)(h_1(t)+z_1)}}_{(1)}$$
$$=\int_0^t \underbrace{e^{(m_0-d)(p_0+\frac{d}{m_0}q_0)(h_1(s)+z_1)}}_{(2)}$$
$$\cdot\left[\underbrace{-\frac{x(r^*)^2}{x'(r^*)}\left(q_0-p_0-2q_0\frac{d}{m_0}\right)}_{(3)}+\underbrace{(m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)x(r^*)\frac{\partial V(d+1,s)}{\partial d}}_{(4)}\right]\mathrm{d}s, \tag{A.21}$$

For all $d > m_0\left(\frac{1}{2}-\frac{p_0}{2q_0}\right)$, we show $\frac{\partial V(d,t)}{\partial d}\le 0$ by induction. When $d=m_0-1$, (4) is zero. The sign of $\frac{\partial V}{\partial d}$ depends on $(1),(2),(3)$. $(1),(2)$ are always nonnegative, and (3) is negative when $d>m_0\left(\frac{1}{2}-\frac{p}{2q_0}\right)$. Suppose $\frac{\partial V(k,t)}{\partial k}\le 0$ for all $m_0\left(\frac{1}{2}-\frac{p_0}{2q_0}\right)< k=d+1,\ldots,m_0-2$. We will then show $\frac{\partial V(d,t)}{\partial d}\le 0$. According to (A.21), $\frac{\partial V(d,t)}{\partial d}\le 0$ because all $(1),(2),(3),(4)\le 0$.

3. Taking the partial integral of (II.3.2) w.r.t r^* and rearranging the terms, we have that

$$\frac{\partial r^*(d,t)}{\partial\left[V(d,t)-V(d+1,t)\right]}\left[\frac{2x'(r^*)^2-x(r^*)x''(r^*)}{x'(r^*)^2}\right]=1. \tag{A.22}$$

Hence, taking the partial derivative of (II.3.3) w.r.t. p_0 and using (A.22), we have

$$\frac{\partial^2 V(d,t)}{\partial p_0\partial t}=-\frac{x(r^*)^2}{x'(r^*)}(m_0-d)$$
$$-(m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)\frac{2x'(r)^2x(r)-x(r)^2x''(r)}{x'(r)^2}\Big|_{r=r^*}$$

$$\cdot \frac{\partial r^*(d,t)}{\partial\left[V(d,t)-V(d+1,t)\right]}\frac{\partial[V(d,t)-V(d+1,t)]}{\partial p_0}.$$

Defining $g(d,t) := \frac{\partial V(d,t)}{\partial p_0}$, we have $g(d,0)=0$ and

$$\begin{aligned}&\frac{\partial g}{\partial t} + (m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)x(r^*)g \\ &= -\frac{x(r^*)^2}{x'(r^*)}(m_0-d) + (m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)x(r^*)\frac{\partial V(d+1,t)}{\partial p_0},\end{aligned} \tag{A.23}$$

which we solve using the same techniques as (A.20), resulting in

$$\begin{aligned}&\frac{\partial V(d,t)}{\partial p_0}\underbrace{e^{(m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)(h_1(t)+z_1)}}_{(1)} \\ &= \int_0^t \underbrace{e^{(m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)(h_1(s)+z_1)}}_{(2)} \\ &\cdot\left[\underbrace{-\frac{x(r^*)^2}{x'(r^*)}(m_0-d)}_{(3)} + \underbrace{(m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)x(r^*)\frac{\partial V(d+1,s)}{\partial p}}_{(4)}\right]\mathrm{d}s.\end{aligned}$$

As we know $\frac{\partial V(m_0,t)}{\partial p_0} = 0$ for all $t\in[0,T]$, then $\frac{\partial V(m_0-1,t)}{\partial p_0}\geq 0$ for all $t\in[0,T]$ because $(1),(2),(3)$ above are always positive and $(4)=0$. Similarly, we can deduce that $\frac{\partial V(m_0-2,t)}{\partial p_0}, \frac{\partial V(m_0-3,t)}{\partial p_0},\dots,\frac{\partial V(0,t)}{\partial p_0}$ are all nonnegative for all t because (4) in these cases become nonnegative. This proves that $V(d,t)$ is monotone increasing in p_0. We can use the same technique to prove monotonicity of the value function in q_0 and in m_0. Defining $g_1 := \frac{\partial V(d,t)}{\partial q_0}$ and $g_2 := \frac{\partial V(d,t)}{\partial m_0}$ (here we treat m_0 as a continuous parameter) results in the following corresponding ordinary differential equations:

$$\begin{aligned}&\frac{\partial g_1}{\partial t} + (m_0-d)(p_0+\frac{d}{m_0}q_0)x(r^*)g_1 \\ &= -\frac{x(r^*)^2}{x'(r^*)}\underbrace{(m_0-d)\frac{d}{m_0}}_{(5)} + (m_0-d)(p_0+\frac{d}{m_0}q_0)x(r^*)\frac{\partial V(d+1,t)}{\partial q_0}, \\ &\frac{\partial g_2}{\partial t} + (m_0-d)(p_0+\frac{d}{m_0}q_0)x(r^*)g_2 \\ &= -\frac{x(r^*)^2}{x'(r^*)}\underbrace{(p_0+\frac{d^2}{{m_0}^2}q_0)}_{(6)} + (m_0-d)(p_0+\frac{d}{m_0}q_0)x(r^*)\frac{\partial V(d+1,t)}{\partial m_0}.\end{aligned}$$

Note that the difference between the two ODEs are from (5) and (6), which are both positive, so can be analyzed in the same way that we did for (A.23). Therefore, $V(d,t)$ is partially monotone increasing in p_0, q_0, m_0.

□

A.2.5 Proof of Corollary II.1

Proof. For the case where $x(r) = e^{-r}$, (II.3.2) results in $r^*(d,t) = 1 - V(d+1,t) + V(d,t)$. Hence, condition (II.3.3) becomes

$$\frac{\partial}{\partial t} V(d,t) = (m_0 - d)\left(p_0 + \frac{d}{m_0} q_0\right) \frac{e^{-2r^*(d,t)}}{e^{-r^*(d,t)}} = \xi(d) e^{V(d+1,t) - V(d,t) - 1}. \tag{A.24}$$

Given the boundary conditions of (II.3.3), we can solve the system backwards:

- From (A.24) when $d = m_0 - 1$, and since $V(m_0, t) = 0$ for all $t \geq 0$, we know that

$$\frac{\partial}{\partial t} V(m_0 - 1, t) = \xi(m_0 - 1) e^{-1 - V(m_0 - 1, t)}.$$

This partial differential equation is solved by $V(m_0 - 1, t) = \ln\left(\frac{\xi(m_0-1)}{e} t + 1\right)$, hence

$$\frac{\partial V(m_0 - 1, t)}{\partial t} = \left(\frac{\xi(m_0 - 1)}{e}\right) \Big/ \left(\frac{\xi(m_0 - 1)}{e} t + 1\right) \tag{A.25}$$

- From (A.24), and (A.25), we know

$$\frac{\partial V(m_0 - 1, t)}{\partial t} \frac{\partial V(m_0 - 2, t)}{\partial t} = \frac{\xi(m_0 - 2)}{e} \left(\frac{\xi(m_0 - 1)}{e} t + 1\right) e^{-V(m_0 - 2, t) - 2}.$$

This is solved by $V(m_0 - 2, t) = \ln\left(\frac{\xi(m_0-1)\xi(m_0-2)}{2e^2} t^2 + \frac{\xi(m_0-2)}{e} t + 1\right)$. Hence,

$$\frac{\partial V(m_0 - 2, t)}{\partial t} \tag{A.26}$$
$$= \left(\frac{\xi(m_0 - 1)\xi(m_0 - 2)}{e^2} t + \frac{\xi(m_0 - 2)}{e}\right) \Big/ \left(\frac{\xi(m_0 - 1)\xi(m_0 - 2)}{2e^2} t^2 + \frac{\xi(m_0 - 2)}{e} t + 1\right). \tag{A.27}$$

- From (A.24), we know

$$\frac{\partial V(m_0 - 1, t)}{\partial t} \frac{\partial V(m_0 - 2, t)}{\partial t} \frac{\partial V(m_0 - 3, t)}{\partial t} = \xi(m_0 - 1)\xi(m_0 - 2)\xi(m_0 - 3) e^{-V(m_0 - 3, t) - 3}.$$

Substituting (A.25)–(A.27), this reduces to a partial differential equation whose solution is

$$
\begin{aligned}
&V(m_0-3,t)\\
&=\ln\left(\frac{\xi(m_0-1)\xi(m_0-2)\xi(m_0-3)}{3!e^3}t^3+\frac{\xi(m_0-2)\xi(m_0-3)}{2!e^2}t^2+\frac{\xi(m_0-3)}{e}t+1\right).
\end{aligned}
$$

This then provides us with $\partial V(m_0-3,t)/\partial t$.

- We can continue to solve for $V(0,t)$:

$$
\begin{aligned}
&V(0,t)\\
&=\ln\left(\frac{\xi(m_0-1)\xi(m_0-2)...\xi(0)}{m_0!}\left(\frac{t}{e}\right)^{m_0}+\frac{\xi(m_0-2)\xi(m_0-3)...\xi(0)}{(m_0-1)!}\left(\frac{t}{e}\right)^{m_0-1}+...+1\right)\\
&=\ln\left(\sum_{j=1}^{m_0}\frac{\prod_{i=0}^{j-1}\xi(i)}{j!}\left(\frac{t}{e}\right)^{j}+1\right).
\end{aligned}
$$

□

A.2.6 Proof of Proposition II.2

Proof. It suffices to show the Hessian matrix of $\sum_{i=0}^{\widehat{D}_{t-}}\ln f_i(\beta)$ with respect to β is negative definite. Note that for $i=0,1,2,3,...,\widehat{D}_{t-}-1$,

$$
\nabla_\beta^2\ln f_i(\beta)=\frac{-1}{(\beta_1+\beta_2 i+\beta_3 i^2)^2}\cdot\begin{pmatrix}1 & i & i^2\\ i & i^2 & i^3\\ i^2 & i^3 & i^4\end{pmatrix},
$$

and $\nabla_\beta^2\ln f_i(\beta)=0$ for $i=\widehat{D}_{t-}$. Hence, for any $\mathbf{z}=(z_1,z_2,z_3)^\top$,

$$
\mathbf{z}^\top\nabla_\beta^2\left(\sum_{i=0}^{\widehat{D}_{t-}}\ln f_i(\beta)\right)\mathbf{z}=-\sum_{i=0}^{\widehat{D}_{t-}-1}\frac{(z_1+iz_2+i^2z_3)^2}{(\beta_1+\beta_2 i+\beta_3 i^2)^2}\le 0
$$

Hence, the Hessian matrix is negative semidefinite.

To show that the Hessian matrix is negative definite, we need the additional condition that $\widehat{D}_{t-}\ge 3$. Under this condition, for any $\mathbf{z}\neq\mathbf{0}$,

$$
\mathbf{z}^\top\nabla_\beta^2\left(\sum_{i=0}^{\widehat{D}_{t-}}\ln f_i(\beta)\right)\mathbf{z}=-\frac{z_1^2}{\beta_1^2}-\frac{(z_1+z_2+z_3)^2}{(\beta_1+\beta_2+\beta_3)^2}-\frac{(z_1+2z_2+4z_3)^2}{(\beta_1+2\beta_2+4\beta_3)^2}-\sum_{i=3}^{\widehat{D}_{t-}-1}\frac{(z_1+iz_2+i^2z_3)^2}{(\beta_1+\beta_2 i+\beta_3 i^2)^2}.
$$

Note that $z_1 = 0$, $z_1 + z_2 + z_3 = 0$ and $z_1 + 2z_2 + 4z_3 = 0$ can only occur simultaneously if $\mathbf{z} = \mathbf{0}$. Hence, $-\frac{z_1^2}{\beta_1^2} - \frac{(z_1+z_2+z_3)^2}{(\beta_1+\beta_2+\beta_3)^2} - \frac{(z_1+2z_2+4z_3)^2}{(\beta_1+2\beta_2+4\beta_3)^2}$ is strictly less than zero for any $\mathbf{z} \neq \mathbf{0}$. This means that we need the condition that $\widehat{D}_{t-} \geq 3$ for $\nabla^2_\beta(\sum_{i=0}^{\widehat{D}_{t-}} \ln f_i(\beta)) \prec 0$. Since $\mathcal{L}_t(\widehat{\mathbf{U}}_t|\ \beta) = \sum_{i=0}^{\widehat{D}_{t-}} \ln f_i(\beta)$, we can conclude that $\nabla^2_\beta \mathcal{L}_t(\widehat{\mathbf{U}}_t|\ \beta) \prec 0$ when $\widehat{D}_{t-} \geq 3$. $\square$

A.2.7 Proof of Lemma II.3

Proof. All the expectations in this proof are conditioning on $D_t^\pi = k$ where $k \geq 3$. For simplicity of notation, we will use D_t instead of D_t^π to denote the cumulative adoptions at time t. Since $D_t \geq 3$, we know that the ML estimator $\hat{\theta}_t$ is unique.

Note from (II.4.2) that if either $\hat{p}_t = +\infty$ or $\hat{q}_t = +\infty$ or $\hat{m}_t = +\infty$, then the likelihood function is 0. Then, we know there exist finite $\bar{\delta}_1, \bar{\delta}_2, \bar{\delta}_3$ such that $\hat{p}_t \leq p_0 + \bar{\delta}_1$, $\hat{q}_t \leq q_0 + \bar{\delta}_2$ and $\hat{m}_t \leq m_0 + \bar{\delta}_3$. Note that the ML estimator $\hat{\theta}_t = (\hat{p}_t, \hat{q}_t, \hat{m}_t)$ can be written as

$$\hat{\theta}_t = \arg\max_{\theta \geq 0} \mathcal{L}_t(\widehat{\mathbf{U}}_t; \theta) = \theta_0 + \arg\min_{u \geq -\theta_0} -\sum_{i=0}^{D_t} \ln \frac{f_i(\theta_0 + u)}{f_i(\theta_0)},$$

where $u = (u_p, u_q, u_m)$, $\theta_0 = (p_0, q_0, m_0)$, and $f_i(\theta)$ is defined in (II.4.2). If we denote the optimizer of the right-hand side as $\hat{u} = (\hat{u}_p, \hat{u}_q, \hat{u}_m)$, then $\hat{\theta}_t = \theta_0 + \hat{u}$.

We analyze the estimation error $|\hat{p}_t - p_0|$. Suppose $|\hat{p}_t - p_0| > \delta$ for some $\bar{\delta}_1 \geq \delta > 0$. This implies that $\hat{u}_p$ lies outside $[-\delta, \delta]$. Since the objective function on the right-hand-side is 0 when $u = 0$, and since the log-likelihood function is continuous and element-wise concave in p, then either

$$-\sum_{i=0}^{D_t} \ln \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)} \leq 0 \quad \text{or} \quad -\sum_{i=0}^{D_t} \ln \frac{f_i(\theta_0 - \delta e_1)}{f_i(\theta_0)} \leq 0,$$

where $e_1 := (1, 0, 0)$. Note that under the Markovian Bass model, the value $f_i(\theta)$ for any θ is stochastic since its value depends on t_i and t_{i+1}, which are random adoption times. Here, t_i denotes the time of the i-th adoption, where $i = 0, \ldots, D_t$.

Let $\mathbb{P}_{\theta_0}(\cdot)$ denote the probability under a demand process that follows a Markovian Bass model with parameter vector $\theta_0 = (p_0, q_0, m_0)$. Therefore,

$$\begin{aligned}
&\mathbb{P}_{\theta_0}\{|\hat{p}_t - p_0| > \delta\} \\
&\leq \mathbb{P}_{\theta_0}\left\{-\sum_{i=0}^{D_t} \ln \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)} \leq 0\right\} + \mathbb{P}_{\theta_0}\left\{-\sum_{i=0}^{D_t} \ln \frac{f_i(\theta_0 - \delta e_1)}{f_i(\theta_0)} \leq 0\right\}
\end{aligned}$$

$$
\begin{aligned}
&\leq 2\mathbb{P}_{\theta_0}\left\{-\sum_{i=0}^{D_t}\ln\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}\leq 0\right\} = 2\mathbb{P}_{\theta_0}\left\{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}\geq 1\right\}\\
&= 2\mathbb{P}_{\theta_0}\left\{\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\geq 1\right\}\leq 2\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\right)\\
&= 2\mathbb{E}_{\theta_0}\left(\mathbb{E}_{\theta_0}\left(\cdots\mathbb{E}_{\theta_0}\left(\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)\mid\mathcal{F}_{t_{D_t-2}}\right)\cdots\mid\mathcal{F}_{t_1}\right)\mid\mathcal{F}_0\right).
\end{aligned}
\tag{A.28}
$$

The second inequality is because f_i is an increasing function in p. The last equality is due to the law of iterated expectations.

We next analyze (A.28) starting from the innermost conditional expectation. We have

$$
\begin{aligned}
&\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\mathbb{E}_{\theta_0}\left(\sqrt{\frac{f_{D_t}(\theta_0+\delta e_1)}{f_{D_t}(\theta_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)\\
&=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\left(\int_{t_{D_t-1}}^{\infty}\sqrt{\frac{f_{D_t}(\theta_0+\delta e_1)}{f_{D_t}(\theta_0)}}f_{D_t}(\theta_0)\mathrm{d}t_{D_t}\right)\\
&=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\left(\int_{t_{D_t-1}}^{\infty}\sqrt{f_{D_t}(\theta_0+\delta e_1)}\sqrt{f_{D_t}(\theta_0)}\mathrm{d}t_{D_t}\right).
\end{aligned}
\tag{A.29}
$$

The first equality is because $\{f_i(\theta), i=0,\ldots,D_t-1\}$ are all $\mathcal{F}_{t_{D_t-1}}$-measurable. The second equality is because, given the information set $\mathcal{F}_{t_{D_t-1}}$, $f_{D_t}(\theta_0)$ is the conditional probability distribution of the adoption time t_{D_t} under a Markovian Bass model with parameter θ_0. Hence, we next want to derive a bound on $\int_{t_{D_t-1}}^{\infty}\sqrt{f_{D_t}(\theta_0+\delta e_1)}\sqrt{f_{D_t}(\theta_0)}\mathrm{d}t_{D_t}$.

Note that

$$
\begin{aligned}
&\frac{1}{2}\int_{t_{D_t-1}}^{\infty}\left(\sqrt{f_{D_t}(\theta_0+\delta e_1)}-\sqrt{f_{D_t}(\theta_0)}\right)^2\mathrm{d}t_{D_t}\\
&=\frac{1}{2}\int_{t_{D_t-1}}^{\infty}\left(f_{D_t}(\theta_0+\delta e_1)+f_{D_t}(\theta_0)-2\sqrt{f_{D_t}(\theta_0+\delta e_1)f_{D_t}(\theta_0)}\right)\mathrm{d}t_{D_t}\\
&=1-\int_{t_{D_t-1}}^{\infty}\sqrt{f_{D_t}(\theta_0+\delta e_1)f_{D_t}(\theta_0)}\mathrm{d}t_{D_t},
\end{aligned}
$$

where the last equality is because the integral of the probability density function

$$\int_{t_{D_{t-1}}}^{\infty} f_{D_t}(\theta) \mathrm{d}t_{D_t}$$

is equal to 1 for any θ. Therefore,

$$\int_{t_{D_{t-1}}}^{\infty} \sqrt{f_{D_t}(\theta_0 + \delta e_1) f_{D_t}(\theta_0)} \mathrm{d}t_{D_t} = 1 - \frac{1}{2} \int_{t_{D_{t-1}}}^{\infty} \left(\sqrt{f_{D_t}(\theta_0 + \delta e_1)} - \sqrt{f_{D_t}(\theta_0)} \right)^2 \mathrm{d}t_{D_t}. \tag{A.30}$$

The integral on the right-hand side is the Hellinger distance between $f_{D_t}(\theta_0 + \delta e_1)$ and $f_{D_t}(\theta_0)$, which are probability densities of the adoption time t_{D_t}.

Note that the Hellinger distance can be lower bounded by the K-L divergence (corollary 4.9 in *Taneja and Kumar* 2004). Specifically,

$$\frac{1}{2} \int_{t_{D_{t-1}}}^{\infty} \left(\sqrt{f_{D_t}(\theta_0 + \delta e_1)} - \sqrt{f_{D_t}(\theta_0)} \right)^2 \mathrm{d}t_{D_t} \geq \frac{1}{4\sqrt{R}} \mathbb{E}_{\theta_0} \left(\ln \frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0 + \delta e_1)} \mid \mathcal{F}_{t_{D_{t-1}}} \right), \tag{A.31}$$

where R is a constant such that $R \geq \min_\delta \frac{1}{f_{D_t}(\theta_0 + \delta e_1)} \geq \frac{1}{m_0 p_0}$. We will next bound the right-hand side of (A.31).

Define $C_I := (p_0 + \bar{\delta}_1 + q_0)^2$. Note that

$$\frac{\partial^2}{\partial \delta^2} \ln \frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0 + \delta e_1)} = \frac{1}{(p_0 + \delta + \frac{D_t}{m_0} q_0)^2} \geq \frac{1}{(p_0 + \bar{\delta}_1 + q_0)^2} = \frac{1}{C_I},$$

where the inequality is because $p_0 + \delta \leq p_0 + \bar{\delta}_1$.

Furthermore, since the expectation of the Fisher score under the true parameter is zero, we have

$$\mathbb{E}_{\theta_0} \left(\frac{\partial}{\partial \delta} \ln \frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0 + \delta e_1)} \bigg|_{\delta = 0} \mid \mathcal{F}_{t_{D_{t-1}}} \right) = 0.$$

Hence, we have

$$\mathbb{E}_{\theta_0}\left(\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0+\delta e_1)}\mid\mathcal{F}_{t_{D_t-1}}\right)=\mathbb{E}_{\theta_0}\left(\int_0^\delta\frac{\partial}{\partial z}\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0+ze_1)}\mathrm{d}z\mid\mathcal{F}_{t_{D_t-1}}\right)$$
$$=\mathbb{E}_{\theta_0}\left(\int_0^\delta\left(\frac{\partial}{\partial z}\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0+ze_1)}-\frac{\partial}{\partial z}\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0+ze_1)}\Big|_{z=0}\right)\mathrm{d}z\mid\mathcal{F}_{t_{D_t-1}}\right)$$
$$=\mathbb{E}_{\theta_0}\left(\int_0^\delta\int_0^z\frac{\partial^2}{\partial z'^2}\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0+z'e_1)}\mathrm{d}z'\mid\mathcal{F}_{t_{D_t-1}}\right)\geq\frac{1}{2C_I}\delta^2.$$

Therefore, (A.31) reduces to

$$\frac{1}{4\sqrt{R}C_I}\delta^2\leq\int_{t_{D_t-1}}^\infty\left(\sqrt{f_{D_t}(\theta_0+\delta e_1)}-\sqrt{f_{D_t}(\theta_0)}\right)^2\mathrm{d}t_{D_t}. \tag{A.32}$$

Hence, from (A.30), we have

$$\int_{t_{D_t-1}}^\infty\sqrt{f_{D_t}(\theta_0+\delta e_1)f_{D_t}(\theta_0)}\mathrm{d}t_{D_t}=1-\frac{1}{2}\int_{t_{D_t-1}}^\infty\left(\sqrt{f_{D_t}(\theta_0+\delta e_1)}-\sqrt{f_{D_t}(\theta_0)}\right)^2\mathrm{d}t_{D_t}$$
$$\leq\exp\left(-\frac{1}{2}\int_{t_{D_t-1}}^\infty\left(\sqrt{f_{D_t}(\theta_0+\delta e_1)}-\sqrt{f_{D_t}(\theta_0)}\right)^2\mathrm{d}t_{D_t}\right)\leq\exp\left(-\frac{1}{8\sqrt{R}C_I}\delta^2\right),$$

where the first inequality is because $e^{-x}\geq 1-x$ for all x. The second inequality is from (A.32). Hence, from (A.29), we have

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)\leq\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\cdot\exp\left(-\frac{1}{8\sqrt{R}C_I}\delta^2\right). \tag{A.33}$$

This provides a bound for the innermost conditional expectation in (A.28).

Observe that all the terms in the bound (A.33) are $\mathcal{F}_{t_{D_t-2}}$-measurable, except for the term $\sqrt{f_{D_t-1}(\theta_0+\delta e_1)/f_{D_t-1}(\theta_0)}$. Taking the conditional expectation of both sides in (A.33) given $\mathcal{F}_{t_{D_t-2}}$, and using the same logic as the above arguments to bound the right-hand side, we have

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\mid \mathcal{F}_{t_{D_t-2}}\right) \leq \sqrt{\prod_{i=0}^{D_t-2}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\cdot \exp\left(-\frac{2}{8\sqrt{R}C_I}\delta^2\right)$$

We can proceed iteratively to evaluate (A.28) as we take conditional expectations given $\mathcal{F}_{t_{D_t}-3}$, $\mathcal{F}_{t_{D_t}-4}$, $\mathcal{F}_0$, resulting in

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\right) \leq \mathbb{E}_{\theta_0}\left(\exp\left(-\frac{D_t+1}{8\sqrt{R}C_I}\delta^2\right)\right)$$

Hence, we have that

$$\mathbb{P}_{\theta_0}\{|\hat{p}_t-p_0|>\delta\mid D_t=k\} \leq 2\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\mid D_t=k\right) \leq 2\exp\left(-\frac{k+1}{8\sqrt{R}C_I}\delta^2\right)$$

if $\delta \leq \bar{\delta}_1$ and otherwise, $\mathbb{P}_{\theta_0}\{|\hat{p}_t-p_0|>\delta\mid D_t=k\}=0$. This implies that

$$\begin{aligned}
&\mathbb{E}_{\theta_0}\left[(\hat{p}_t-p_0)^2\mid D_t=k\right] \\
&= \int_0^\infty \mathbb{P}_{\theta_0}\left\{(\hat{p}_t-p_0)^2>\delta\mid D_t=k\right\}\mathrm{d}\delta = \int_0^\infty \mathbb{P}_{\theta_0}\left\{|\hat{p}_t-p_0|^2>\sqrt{\delta}\mid D_t=k\right\}\mathrm{d}\delta \\
&\leq \int_0^\infty 2\exp\left(-\frac{k+1}{16\sqrt{R}C_I}\delta\right)\mathrm{d}\delta = \frac{8\sqrt{R}C_I}{k+1}.
\end{aligned}$$

Thus, we have shown that the mean squared error of $\hat{p}_t$ is bounded by $\frac{\alpha_p}{k+1}$ for some α_p that is independent of m_0 and t. Hence, to prove the lemma, we only need to show a similar bound for $\hat{m}_t, \hat{q}_t$. Similar bounds can be obtained for $\hat{m}_t, \hat{q}_t$ following the same steps with the only difference on the definition of C_I.

For $\hat{q}_t$, the estimation variance of $\hat{q}_t$ grows as D_t/m_0 approaches zero. To avoid this issue when D_t/m_0 is small, we perform a transformation on the parameters of the likelihood function. Specifically, we let $p'=p-q$. Thus, MLE estimates the model parameters $\theta'=(p',q,m)$ of a Markovian Bass model where the adoption rate is $\lambda(j,r;\theta')=(m-j)\left(p'+q\left(1+\frac{j}{m}\right)\right)x(r)$. Note that the analysis of the estimation error for $\hat{p}'_t$ is the same as that for $\hat{p}_t$. With the transformation, we can safely write the second order derivative of the log-likelihood function with respect to q. We have

$$\mathbb{E}_{\theta_0'}\left[\frac{\partial^2}{\partial\delta^2}\ln\frac{f_{D_t}(\theta_0')}{f_{D_t}(\theta_0'+\delta e_2)}\mid\mathcal{F}_{t_{D_t-1}}\right]=\mathbb{E}_{\theta_0'}\left[\frac{\left(1+\frac{D_t}{m_0}\right)^2}{\left({p_0}'+\left(1+\frac{D_t}{m_0}\right)(q_0+\delta)\right)^2}\mid\mathcal{F}_{t_{D_t-1}}\right]$$
$$\geq\mathbb{E}_{\theta_0'}\left[\frac{\left(1+\frac{D_t}{m_0}\right)^2}{\left({p_0}'+\left(1+\frac{D_t}{m_0}\right)(q_0+\bar{\delta}_2)\right)^2}\mid\mathcal{F}_{t_{D_t-1}}\right],$$

where the inequality is because $q_0+\delta\leq q_0+\bar{\delta}_2$. Defining $C_I:=(p_0+q_0+\bar{\delta}_2)^2$, we have that

$$\mathbb{E}_{\theta_0'}\left[\frac{\left(1+\frac{D_t}{m_0}\right)^2}{\left(p_0+\left(1+\frac{D_t}{m_0}\right)(q_0+\bar{\delta}_2)\right)^2}\mid\mathcal{F}_{t_{D_t-1}}\right]\geq\frac{1}{\left(p_0+q_0+\bar{\delta}_2\right)^2}=\frac{1}{C_I}.$$

Following the same steps in bounding the estimation error of $\hat{p}_t$, we know

$$\mathbb{P}_{\theta_0}\{|\hat{q}_t-q_0|>\delta\mid D_t=k\}\leq 2\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=1}^{D_t}\frac{f_i(\theta_0+\delta e_2)}{f_i(\theta_0)}}\Big|D_t=k\right)\leq 2\exp\left(-\frac{k+1}{8\sqrt{R}C_I}\delta^2\right).$$

This implies that

$$\begin{aligned}&\mathbb{E}_{\theta_0}\left[(\hat{q}_t-q_0)^2\mid D_t=k\right]\\&=\int_0^\infty\mathbb{P}_{\theta_0}\left\{(\hat{q}_t-q_0)^2>\delta\mid D_t=k\right\}\mathrm{d}\delta=\int_0^\infty\mathbb{P}_{\theta_0}\left\{|\hat{q}_t-q_0|^2>\sqrt{\delta}\mid D_t=k\right\}\mathrm{d}\delta\\&\leq\int_0^\infty 2\exp\left(-\frac{k+1}{16\sqrt{R}C_I}\delta\right)\mathrm{d}\delta=\frac{8\sqrt{R}C_I}{k+1}.\end{aligned}\tag{A.34}$$

Thus, we have shown that the mean squared error of $\hat{q}_t$ is bounded by $\frac{\alpha_q}{k+1}$ for some α_q that is independent of m_0 and t.

For $\hat{m}_t$, we have

$$\frac{\partial^2}{\partial\delta^2}\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0+\delta e_3)}\geq\frac{1}{(m_0+\delta-D_t)^2}\geq\frac{1}{(\bar{\delta}_3)^2}.$$

So C_I for $\hat{m}_t$ is $(\bar{\delta}_3)^2$. Following the same steps for bounding the mean squared estimation error of $\hat{q}_t$ and $\hat{p}_t$, this means that the mean squared error of $\hat{m}_t$ is bounded by $\alpha_m/(k+1)$ for some α_m that is independent of m_0 and t. □

A.2.8 Lemma A.2 and proof

We next state a result that is useful for the proofs of Proposition II.3 and Lemma II.4.

Lemma A.2. Given any two pricing sample paths $r = (r_t, t \geq 0)$ and $r' = (r_t, t \geq 0)$, if $D^{r,m_0} = (D_t^{r,m_0}, t \geq 0)$ and $D^{r',m_0} = (D_t^{r',m_0}, t \geq 0)$, respectively, denote the cumulative adoption process with market potential m_0, then for any $t \geq 0$,

$$\mathbb{E}_{\theta_0}\left|\frac{D_t^{r,m_0}}{m_0} - \frac{D_t^{r',m_0}}{m_0}\right| = \alpha_1 \frac{t}{e^t}\left|r_t - r'_t\right| + \mathcal{O}\left(\frac{1}{\sqrt{m_0}}\right), \tag{A.35}$$

$$\left|\mathbb{E}_{\theta_0}\left(\frac{D_t^{r,m_0}}{m_0} - \frac{D_t^{r',m_0}}{m_0}\right)\right| = \alpha_2 \frac{t}{e^t}\left|r_t - r'_t\right| \tag{A.36}$$

for some $\alpha_1 > 0, \alpha_2 > 0$ independent of m_0.

Observe from Lemma A.2 that the expectation of the absolute difference is greater than the absolute value of the expected difference by $\mathcal{O}\left(1/\sqrt{m_0}\right)$. This is because the *uncertainty* of the Markovian Bass model, Var $(D_t^{r,m_0}/m_0)$, decreases in the order of $\mathcal{O}(1/m_0)$ (Proposition II.1).

Proof. We first define for any $t \geq 0$,

$$F_t^r = \frac{1 - e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}}{1 + q_0/p_0 e^{-(p_0+q_0)\int_0^t x(r_s)\mathrm{d}s}},$$
$$F_t^{r'} = \frac{1 - e^{-(p_0+q_0)\int_0^t x(r'_s)\mathrm{d}s}}{1 + q_0/p_0 e^{-(p_0+q_0)\int_0^t x(r'_s)\mathrm{d}s}},$$

which are the deterministic Bass functions under price processes r and r'. We have

$$\begin{aligned}
\left|F_t^r - F_t^{r'}\right| &= \int_{\int_0^t x(r_s)\mathrm{d}s}^{\int_0^t x(r'_s)\mathrm{d}s} \frac{(1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)X}}{\left(1+q_0/p_0 e^{-(p_0+q_0)X}\right)^2}\mathrm{d}X \\
&\leq \frac{(1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)\int_0^t x(r_u^\xi)\mathrm{d}u}}{\left(1+q_0/p_0 e^{-(p_0+q_0)\int_0^t x(r_u^\xi)\mathrm{d}u}\right)^2}\left|\int_0^t x(r_s)\mathrm{d}s - \int_0^t x(r'_s)\mathrm{d}s\right| \\
&\leq (1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)\bar{x}^l t}\left|\int_0^t x(r_s)\mathrm{d}s - \int_0^t x(r'_s)\mathrm{d}s\right| = \frac{t}{e^t}\mathcal{O}\left(\left|r_t - r'_t\right|\right),
\end{aligned}$$

where $\int_0^t x(r_s^\xi)\mathrm{d}s$ is in the between of $\int_0^t x(r_s)\mathrm{d}s$ and $\int_0^t x(r_s')\mathrm{d}s$. Here the second inequality comes from Assumption II.1 (ii).

Note that for any $t \geq 0$,

$$\mathbb{E}_{\theta_0}\left|\frac{D_t^{r,m_0}}{m_0}-\frac{D_t^{r',m_0}}{m_0}\right| \leq \left|F_t^r - F_t^{r'}\right| + \mathbb{E}_{\theta_0}\left|\frac{D_t^{r,m_0}}{m_0}-F_t^r\right| + \mathbb{E}_{\theta_0}\left|\frac{D_t^{r',m_0}}{m_0}-F_t^{r'}\right|$$
$$= \left|F_t^r - F_t^{r'}\right| + \mathcal{O}\left(\frac{1}{\sqrt{m_0}}\right).$$

where the last relationship follows from Lemma II.2. Using the bound on $|F_t^r - F_t^{r'}|$ proves (A.35).

We next prove (A.36). For any $t \geq 0$, define

$$F_t^{r,m_0} := \mathbb{E}_{\theta_0}\left(\frac{D_t^{r,m_0}}{m_0}\right), \qquad F_t^{r',m_0} := \mathbb{E}_{\theta_0}\left(\frac{D_t^{r',m_0}}{m_0}\right).$$

Following the proof in Lemma II.2 in deriving (A.14), both F_t^{r,m_0} and F_t^{r',m_0} can be expressed in the following form:

$$F_t^{r,m_0} = \frac{1-\exp\left(-(p_0+q_0)\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s\right)}{1+\frac{q_0}{p_0}\exp\left(-(p_0+q_0)\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s\right)}.$$

Note that

$$\frac{\mathrm{d}}{\mathrm{d}X}\left(\frac{1-e^{-(p_0+q_0)X}}{1+\frac{q_0}{p_0}e^{-(p_0+q_0)X}}\right) = \frac{(1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)X}}{(1+q_0/p_0e^{-(p_0+q_0)X})^2}.$$

Therefore,

$$
\begin{aligned}
&\left|\mathbb{E}_{\theta_0}\left(\frac{D_t^{r,m_0}}{m_0}-\frac{D_t^{r',m_0}}{m_0}\right)\right| = |F_t^{r,m_0}-F_t^{r',m_0}| \\
&= \int_{\int_0^t\left(1-\frac{1}{m_0}\frac{q_0F_s^{r,m_0}}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s}^{\int_0^t\left(1-\frac{1}{m_0}\frac{q_0F_s^{r',m_0}}{p_0+q_0F_s^{r',m_0}}\right)x(r'_s)\mathrm{d}s} \frac{(1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)X}}{\left(1+q_0/p_0e^{-(p_0+q_0)X}\right)^2}\mathrm{d}X \\
&\leq \frac{(1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)\bar{x}^l t(1-1/m_0)}}{(1)^2}\left|\int_0^t x(r_s)\mathrm{d}s-\int_0^t x(r'_s)\mathrm{d}s\right| \\
&\leq (1+q_0/p_0)(p_0+q_0)e^{-(p_0+q_0)\bar{x}^l t(1-1/m_0)}\left|\int_0^t x(r_s)\mathrm{d}s-\int_0^t x(r'_s)\mathrm{d}s\right| = \frac{t}{e^t}\mathcal{O}\left(|r_t-r'_t|\right),
\end{aligned}
$$

where the first inequality is from Assumption II.1 (iii). This proves (A.36). □

A.2.9 Lemma A.3 and proof

We next state a result that is also useful for the proof of Proposition II.3.

Lemma A.3. Given any two price sample paths $r=(r_t,t\geq 0)$ and $r'=(r_t,t\geq 0)$, if $D^{r,m_0}=(D_t^{r,m_0},t\geq 0)$ and $D^{r',m_0}=(D_t^{r',m_0},t\geq 0)$, respectively, denote the cumulative adoption process with market potential m_0, then for any $t\geq 0$,

$$
\left|\mathbb{E}_{\theta_0}\left(\frac{\xi(D_t^{r,m_0})}{m_0}-\frac{\xi(D_t^{r',m_0})}{m_0}\right)\right| \leq (p_0+q_0)\alpha_1\frac{t}{e^t}\left|r_t-r'_t\right|, \tag{A.37}
$$

and

$$
\left|\mathbb{E}_{\theta_0}\left(\frac{\xi(D_t^{r,m_0})}{\xi(D_t^{r',m_0})}\right)\right| = 1+\alpha_2\frac{t}{e^t}\frac{|r_t-r'_t|}{m_0} \tag{A.38}
$$

for some $\alpha_1>0,\alpha_2>0$ independent of m_0.

Proof. Using the definition of $\xi(\cdot)$ in (II.2.4), we can write

$$
\frac{\xi(d)}{m_0}=\frac{(m_0-d)}{m_0}\left(p_0+q_0\frac{d}{m_0}\right)=p_0+(q_0-p_0)\left(\frac{d}{m_0}\right)-q_0\left(\frac{d}{m_0}\right)^2.
$$

From this, and using the fact that $\frac{\mathrm{d}}{\mathrm{d}y}\left(p_0+(q_0-p_0)y-q_0y^2\right)=q_0-p_0-2q_0y$, we have

$$\begin{aligned}
\left|\mathbb{E}_{\theta_0}\left(\frac{\xi(D_t^{r,m_0})}{m_0}-\frac{\xi(D_t^{r',m_0})}{m_0}\right)\right| &= \left|\mathbb{E}_{\theta_0}\left(\int\limits_{D_t^{r',m_0}/m_0}^{D_t^{r,m_0}/m_0}(q_0-p_0-2q_0y)\mathrm{d}y\right)\right| \\
&\le \left|\mathbb{E}_{\theta_0}\left(\sup_{y\in[0,1)}(q_0-p_0-2q_0y)\int\limits_{D_t^{r',m_0}/m_0}^{D_t^{r,m_0}/m_0}\mathrm{d}y\right)\right| \\
&\le (p_0+q_0)\left|\mathbb{E}_{\theta_0}\left(\frac{D_t^{r,m_0}}{m_0}-\frac{D_t^{r',m_0}}{m_0}\right)\right|.
\end{aligned}$$

Then, (A.37) follows from Lemma A.2.

To prove (A.38), note that

$$\frac{\mathrm{d}}{\mathrm{d}y}\left(\ln\xi(y)\right)=\frac{q_0-p_0-2q_0y/m_0}{(m_0-y)(p_0+q_0y/m_0)}.$$

Therefore, we have

$$\begin{aligned}
&\left|\mathbb{E}_{\theta_0}\left(\frac{\xi(D_t^{r,m_0})}{\xi(D_t^{r',m_0})}\right)\right| \\
&= \left|\mathbb{E}_{\theta_0}\left(e^{\ln\xi\left(D_t^{r,m_0}\right)-\ln\xi(D_t^{r',m_0})}\right)\right| \\
&= \left|\mathbb{E}_{\theta_0}\left(\exp\left\{\int\limits_{D_t^{r,m_0}}^{D_t^{r',m_0}}\frac{q_0-p_0-2q_0y/m_0}{(m_0-y)(p_0+q_0y/m_0)}\mathrm{d}y\right\}\right)\right| \\
&\le \left|\mathbb{E}_{\theta_0}\left(\exp\left\{\sup_{0\le y\le m_0-1}\frac{q_0-p_0-2q_0y/m_0}{(m_0-y)(p_0+q_0y/m_0)}\int\limits_{D_t^{r,m_0}}^{D_t^{r',m_0}}\mathrm{d}y\right\}\right)\right| \\
&= \left|\mathbb{E}_{\theta_0}\left(\exp\left\{\sup_{0\le y\le m_0-1}\frac{(q_0-p_0)m_0-2q_0y}{m_0(m_0-y)(p_0+q_0y/m_0)}\left(\frac{D_t^{r,m_0}}{m_0}-\frac{D_t^{r',m_0}}{m_0}\right)\right\}\right)\right| \\
&\le \left|\mathbb{E}_{\theta_0}\left(\exp\left\{\sup_{0\le y\le m_0-1}\frac{\max\{q_0-p_0,2q_0\}(m_0-y)}{m_0(m_0-y)(p_0+q_0y/m_0)}\left(\frac{D_t^{r,m_0}}{m_0}-\frac{D_t^{r',m_0}}{m_0}\right)\right\}\right)\right| \\
&\le \left|\mathbb{E}_{\theta_0}\left(\exp\left\{\frac{\max\{q_0-p_0,2q_0\}}{m_0p_0}\left(\frac{D_t^{r,m_0}}{m_0}-\frac{D_t^{r',m_0}}{m_0}\right)\right\}\right)\right| = 1+\frac{t}{e^t}\mathcal{O}\left(\frac{|r_t-r_t'|}{m_0}\right),
\end{aligned}$$

where the second inequality follows from the fact that when $x > y \geq 0$ and $\max\{c_1, c_2\} \geq 0$, $c_1 x - c_2 y \leq \max\{c_1, c_2\}(x - y)$, and the last equality is from Lemma A.2 and

$$\exp\left(\frac{t}{e^t}\mathcal{O}\left(\frac{|r_t - r'_t|}{m_0}\right)\right) = 1 + \frac{t}{e^t}\mathcal{O}\left(\frac{|r_t - r'_t|}{m_0}\right).$$

□

A.2.10 Proof of Proposition II.3

Proof. Recall that $\theta_0 = (p_0, q_0, m_0)$ denotes the true parameter set. To prove the result, we discretize $[0, T]$ into N small intervals with length δt, where δt is arbitrarily small. Let $v(d, n\delta t, \mathcal{F}_t)$ denote the expected revenue-to-go function under policy π when current cumulative demand is d, where $d \in \{0, 1, \ldots, m_0\}$, the remaining time is $n\delta t$, where $n \in \{1, 2, \ldots, N\}$, and the information set is $\mathcal{F}_t$. Note that this expectation is taken with respect to the true parameter set θ_0.

We denote by $r^\pi(d, n\delta t, \mathcal{F}_t)$ the price offered under policy π given the state $(d, n\delta t, \mathcal{F}_t)$. To simplify notation, we will drop $\mathcal{F}_t$ as an argument in v and r^π, but emphasize that the policy π relies on the information set. For any $d \leq m_0 - 1$, we can write the expected revenue-to-go as

$$\begin{aligned} v(d, n\delta t) &= r^\pi(d, n\delta t) \cdot \xi(d) x(r^\pi(d, n\delta t))\delta t \\ &\quad + [v(d+1, (n-1)\delta t) - v(d, (n-1)\delta t)] \cdot \xi(d) x(r^\pi(d, n\delta t))\delta t + v(d, (n-1)\delta t), \end{aligned} \tag{A.39}$$

where the adoption probability $\xi(d)x(r)\delta t$ is under a Markovian Bass demand model with parameter vector θ_0. Note that $v(m_0, n\delta t) = 0$ for any n, since all customers have already adopted.

Given the state $(d, n\delta t)$, where $d \in \{0, \ldots, m_0\}$ and $n \in \{1, \ldots, N\}$, let $V(d, n\delta t)$ be the optimal expected revenue-to function of the hindsight optimal policy π^* which knows the true value θ_0. For any $d \leq m_0 - 1$, $V(d, n\delta t)$ can be expressed as

$$\begin{aligned} V(d, n\delta t) &= r^*(d, n\delta t) \cdot \xi(d) x(r^*(d, n\delta t))\delta t \\ &\quad + [V(d+1, (n-1)\delta t) - V(d, (n-1)\delta t)] \cdot \xi(d) x(r^*(d, n\delta t))\delta t + V(d, (n-1)\delta t), \end{aligned} \tag{A.40}$$

where $r^*(d, t)$ is the optimal price offered under the optimal policy π^* given state (d, t), defined in Theorem II.1. Note that $V(m_0, n\delta t) = 0$ for any n.

Let $D^\pi = (D^\pi_t,\, t \geq 0)$ and $D^* = (D^*_t,\, t \geq 0)$ be the cumulative demand process under π and π^*, respectively. Let $(r^\pi_t,\, t \geq 0)$ and $(r^*_t,\, t \geq 0)$ denote the price process under π and π^*, respectively. For any $n = 0, 1, \ldots, N-1$, we define

$$\begin{aligned}
\Psi_n &:= \mathbb{E}_{\theta_0}\left(|V(D^*_{n\delta t}, T - n\delta t) - v(D^\pi_{n\delta t}, T - n\delta t)| \mid \mathcal{F}_{n\delta t}\right) \\
&= \mathbb{E}_{\theta_0}\left(\left|\sum_{s=n}^{N-1} r^*_{s\delta t} D^*_{s\delta t}\delta t - \sum_{s=n}^{N-1} r^\pi_{s\delta t} D^\pi_{s\delta t}\delta t\right| \mid \mathcal{F}_{n\delta t}\right)
\end{aligned}$$

as the conditional expectation of the difference in the revenue-to-go between π^* and π, starting from time $n\delta t$ on the discretized grid, and given the information available at time $n\delta t$. To prove the proposition, we will use induction to prove for any $n = 0, 1, \ldots, N-1$,

$$\Psi_n = \mathcal{O}\left(\mathbb{E}_{\theta_0}\left[\sum_{s=n}^{N-1} \frac{D^\pi_{s\delta t} + 1}{s\delta t + t_0}(r^\pi_{s\delta t} - r^*_{s\delta t})^2 \delta t \mid \mathcal{F}_{n\delta t}\right]\right). \tag{A.41}$$

Here, $\mathcal{O}$ describes the limiting behavior as m_0 grows, and the terms inside $\mathcal{O}$ are potentially affected by m_0. Proposition II.3 is an implication of this result since $R^* - R(\pi) = \Psi_0$.

To aid in our induction analysis, we next introduce some notation. For a fixed sample path ω, we denote the realization of D^π and D^* as $(d^\pi_{\omega,t},\, t \geq 0)$ and $(d^*_{\omega,t},\, t \geq 0)$, respectively. For a fixed sample size ω, we denote the realization of the price process under π and π^* as $(\rho^\pi_{\omega,t},\, t \geq 0)$ and $(\rho^*_{\omega,t},\, t \geq 0)$, respectively.

Base case: To prove (A.41), we first check the base step at $n = N-1$. In this case, time is $(N-1)\delta t = T - \delta t$, and there is δt time remaining. For a fixed sample size ω, note that

$$\begin{aligned}
&|V(d^*_{\omega,T-\delta t}, \delta t) - v(d^\pi_{\omega,T-\delta t}, \delta t)| \\
&= \left|\rho^*_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^*_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^\pi_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t\right| \\
&\quad \leq \underbrace{\left|\rho^*_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^*_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t\right|}_{\text{(A)}} \\
&\quad\quad + \underbrace{\left|\rho^\pi_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^\pi_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t\right|}_{\text{(B)}}.
\end{aligned} \tag{A.42}$$

We examine (A) and (B) separately.

- **Bounding (A):** Recall that we have proved in Theorem II.1 that the optimal policy maximizes the revenue-to-go for any given state, and satisfies the first order condition for any given state. Therefore, given the state $(d^*_{\omega,T-\delta t}, \delta t)$, we have

$$\frac{\partial[r\xi(d^*_{\omega,T-\delta t})x(r)\delta t]}{\partial r}\Big|_{r=\rho^*_{\omega,T-\delta t}} = 0. \tag{A.43}$$

Then, we can derive the upper bound of (A) as follows:

$$\begin{aligned}
(\mathrm{A}) &= \left| \int_{\rho^\pi_{\omega,T-\delta t}}^{\rho^*_{\omega,T-\delta t}} \frac{\partial[r\xi(d^*_{\omega,T-\delta t})x(r)\delta t]}{\partial r} \mathrm{d}r \right| = \left| \int_{\rho^\pi_{\omega,T-\delta t}}^{\rho^*_{\omega,T-\delta t}} \left(\frac{\partial[r\xi(d^*_{\omega,T-\delta t})x(r)\delta t]}{\partial r} - 0 \right) \mathrm{d}r \right| \\
&= \left| \int_{\rho^\pi_{\omega,T-\delta t}}^{\rho^*_{\omega,T-\delta t}} \left(\frac{\partial[r\xi(d^*_{\omega,T-\delta t})x(r)\delta t]}{\partial r} - \frac{\partial[r\xi(d^*_{\omega,T-\delta t})x(r)\delta t]}{\partial r}\Big|_{r=\rho^*_{\omega,T-\delta t}} \right) \mathrm{d}r \right| \qquad (\mathrm{A.44}) \\
&= \left| \int_{\rho^\pi_{\omega,T-\delta t}}^{\rho^*_{\omega,T-\delta t}} \int_{\rho^*_{\omega,T-\delta t}}^{r} \frac{\partial^2[z\xi(d^*_{\omega,T-\delta t})x(z)\delta t]}{\partial z^2} \mathrm{d}z\mathrm{d}r \right| \\
&\le \xi(d^*_{\omega,T-\delta t})\delta t \cdot \sup_{z\in(-\infty,\infty)} \left| \frac{\partial^2}{\partial z^2}(zx(z)) \right| \cdot \left| \int_{\rho^\pi_{\omega,T-\delta t}}^{\rho^*_{\omega,T-\delta t}} \int_{\rho^*_{\omega,T-\delta t}}^{r} \mathrm{d}z\mathrm{d}r \right| \\
&\le \frac{1}{2} C_{xx}\xi(d^*_{\omega,T-\delta t})(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t})^2 \delta t. \qquad (\mathrm{A.45})
\end{aligned}$$

Here, (A.44) comes from (A.43), while (A.45) comes from Assumption II.1(v). We replace $\xi(d^*_{\omega,T-\delta t})$ in (A.45) by $\frac{\xi(d^*_{\omega,T-\delta t})}{\xi(d^\pi_{\omega,T-\delta t})}\xi(d^\pi_{\omega,T-\delta t})$. Then, according to Lemma A.3, (A.45) is bounded above by

$$\begin{aligned}
&\frac{1}{2} C_{xx}\left(1 + \frac{T-\delta t}{e^{T-\delta t}}\mathcal{O}(1)\right)\xi(d^\pi_{\omega,T-\delta t})(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t})^2\delta t \\
&= \frac{1}{2} C_{xx}\left(1 + \frac{1}{(T-\delta t)^3}o(1)\right)\xi(d^\pi_{\omega,T-\delta t})(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t})^2\delta t. \qquad (\mathrm{A.46})
\end{aligned}$$

- **Bounding (B):** To analyze the upper bound of (B), according to Assumption II.1(v), we know

$$(\mathrm{B}) \le \sup_{r\in(-\infty,\infty)} |rx(r)| \cdot \left|\xi(d^*_{\omega,T-\delta t}) - \xi(d^\pi_{\omega,T-\delta t})\delta t\right| \le C_x \left| \xi(d^*_{\omega,T-\delta t}) \left(\frac{\xi(d^\pi_{\omega,T-\delta t})}{\xi(d^*_{\omega,T-\delta t})} - 1 \right) \delta t \right|. \tag{A.47}$$

We replace $\xi(d^*_{\omega,T-\delta t})$ in (A.47) by $\frac{\xi(d^*_{\omega,T-\delta t})}{\xi(d^\pi_{\omega,T-\delta t})}\xi(d^\pi_{\omega,T-\delta t})$. Then from (A.47) and

Lemma A.3,

$$
\begin{aligned}
&\text{(B)} \qquad \text{(A.48)}\\
&\leq C_x \left| \xi(d^{\pi}_{\omega,T-\delta t}) \left(1 + \frac{T-\delta t}{e^{T-\delta t}} \mathcal{O}\left(\frac{\left|\rho^{\pi}_{\omega,T-\delta t} - \rho^{*}_{\omega,T-\delta t}\right|}{m_0}\right)\right) \frac{T-\delta t}{e^{T-\delta t}} \mathcal{O}\left(\frac{\left|\rho^{\pi}_{\omega,T-\delta t} - \rho^{*}_{\omega,T-\delta t}\right|}{m_0}\right) \delta t \right|\\
&\leq C_x \left| \xi(d^{\pi}_{\omega,T-\delta t}) \left(1 + \frac{1}{(T-\delta t)^3} \mathcal{O}\left(\frac{\left|\rho^{\pi}_{\omega,T-\delta t} - \rho^{*}_{\omega,T-\delta t}\right|}{m_0}\right)\right) \frac{T-\delta t}{e^{T-\delta t}} \mathcal{O}\left(\frac{\left|\rho^{\pi}_{\omega,T-\delta t} - \rho^{*}_{\omega,T-\delta t}\right|}{m_0}\right) \delta t \right|. \qquad \text{(A.49)}
\end{aligned}
$$

Note that the bound on (A) and (B) both rely on the term $\xi(d^{\pi}_{\omega,T-\delta t})$. Therefore, to proceed with the proof, we need the following claim.

Claim A.1. If at time t, the cumulative demand under π is D^{π}_t, the following holds:

$$
\mathbb{E}(\xi(D^{\pi}_t) \mid \mathcal{F}_t) \leq \alpha_1 \left(\frac{\mathbb{E}\left[\int_0^{t+t_0} \xi(D^{\pi}_s) x(r^{\pi}_s) \mathrm{d}s \mid \mathcal{F}_t\right]}{t+t_0} \right) = \alpha_2 \left(\frac{D^{\pi}_t + 1}{t+t_0} \right) \qquad \text{(A.50)}
$$

for some $\alpha_1 > 0, \alpha_2 > 0$ independent of m_0.

To prove the claim, we first notice that for all $j = 0, 1, \ldots, m_0 - 1$ and any $0 < h \leq 1/m_0$, we have $p_0 \leq \xi(j) \leq m_0 \frac{(p_0+q_0)^2}{4q_0}$, which implies $\xi(D^{\pi}_t) \leq \Theta(m_0)$ almost surely.

Since $\xi(d) = (m_0 - d)\left(p_0 + q_0 \frac{d}{m_0}\right)$ is a concave function in d, then we have that, for any $0 \leq s \leq t$,

$$
\xi(D^{\pi}_s) \geq \min\{\xi(0), \xi(D^{\pi}_t)\} = \min\{m_0 p_0, \xi(D^{\pi}_t)\}.
$$

Therefore,

$$
\int_0^t \xi(D^{\pi}_s) x(r^{\pi}_s) \mathrm{d}s \geq \min\{m_0 p_0, \xi(D^{\pi}_t)\} \int_0^t x(r^{\pi}_s) \mathrm{d}s \geq \alpha_3(\xi(D^{\pi}_t)) \int_0^t x(r^{\pi}_s) \mathrm{d}s \geq \alpha_4(\xi(D^{\pi}_t) t)
$$

for some $\alpha_3 > 0, \alpha_4 > 0$ independent of m_0. Here the last inequality comes from Assumption II.1(ii). Then we can take $\mathbb{E}(\cdot \mid \mathcal{F}_t)$ on both sides and yields

$$
\mathbb{E}(\xi(D^{\pi}_t) \mid \mathcal{F}_t) \leq \alpha_1 \left(\frac{\mathbb{E}\left[\int_0^t \xi(D^{\pi}_s) x(r^{\pi}_s) \mathrm{d}s \mid \mathcal{F}_t\right]}{t} \right)
$$

with $\alpha_1 = 1/\alpha_4$. which gives us (A.50). Note that we use $D^{\pi}_t + 1$ and $t + t_0$ in the final bound to avoid meaningless fractions. This concludes the claim.

Now we are ready to prove the base case. From (A.46) and (A.49), we have that the following constraint holds almost surely:

$$
\begin{aligned}
&|V(D^*_{T-\delta t},\delta t) - v(D^\pi_{T-\delta t},\delta t)| \\
&\quad \le \xi(D^\pi_{T-\delta t}) \left(1 + \frac{1}{(T-\delta t)^3} o\left(\frac{\left|\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t}\right|}{m_0}\right)\right)\delta t \qquad \text{(A.51)}\\
&\quad \cdot \left[\frac{1}{2}C_{xx}\left(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t}\right)^2 + C_x \frac{T-\delta t}{e^{T-\delta t}} \mathcal{O}\left(\frac{\left|\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t}\right|}{m_0}\right)\right]. \qquad \text{(A.52)}
\end{aligned}
$$

From the condition of the proposition, $\mathbb{E}\left[\left|\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t}\right|\right] \ge \alpha\left(\frac{T-\delta t}{m_0 e^{T-\delta t}}\right)$. Therefore, taking the conditional expectation of (A.52) given $\mathcal{F}_{T-\delta t}$, (A.46) dominates (A.49). This results in

$$
\begin{aligned}
\Psi_{N-1} &= \mathcal{O}(1)\mathbb{E}\left(\xi(D^\pi_{T-\delta t}) \mid \mathcal{F}_{T-\delta t}\right)\left(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t}\right)^2 \delta t \\
&\le \mathcal{O}(1)\frac{D^\pi_{T-\delta t}+1}{T-\delta t + t_0}\left(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t}\right)^2 \delta t.
\end{aligned}
$$

The last relation is due to Claim A.1. Here, $t_0 = \Omega({m_0}^{-1})$, which can be interpreted as the inter-arrival time to have one more adoption. It is at least in the order of ${m_0}^{-1}$ because the expected adoption rate is linear in $\xi(j), j = 0, 1, \ldots, m_0 - 1$, and $\xi(j)$ is always less than $m_0 \frac{(p_0+q_0)^2}{4q_0}$. This finishes our base step.

Inductive step: We assume that the result (A.41) holds for $n+1$. Specifically,

$$
\begin{aligned}
\Psi_{n+1} &:= \mathbb{E}\left(\left|V(D^*_{(n+1)\delta t}, T-(n+1)\delta t) - v(D^\pi_{(n+1)\delta t}, T-(n+1)\delta t)\right| \mid \mathcal{F}_{(n+1)\delta t}\right) \\
&= \mathcal{O}\left(\mathbb{E}_{\theta_0}\left[\sum_{s=n+1}^{N-1} \frac{D^\pi_{s\delta t}+1}{s\delta t + t_0}(\rho^\pi_{\omega,s\delta t} - \rho^*_{\omega,s\delta t})^2 \delta t \mid \mathcal{F}_{(n+1)\delta t}\right]\right) \qquad \text{(A.53)}
\end{aligned}
$$

where the $\mathcal{O}$ represents the limiting effect of increasing m_0. We will prove that this implies that it also holds for n.

For a fixed sample ω, we have that

$$
\begin{aligned}
&|V\left(d^*_{\omega,n\delta t}, T-n\delta t\right) - v\left(d^\pi_{\omega,n\delta t}, T-n\delta t\right)| \\
&= \Big|\, \rho^*_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t - \rho^\pi_{\omega,n\delta t}\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
&\quad + \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t \\
&\quad - \left[v(d^\pi_{\omega,n\delta t}+1, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
&\quad + V(d^*_{\omega,n\delta t}, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)\,\Big|
\end{aligned}
$$

$$\leq \text{(A)} + \text{(B)} + \text{(C)} \tag{A.54}$$

where $\text{(C)} = |V(d^*_{\omega,n\delta t}, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)|$,

$$\begin{aligned}
\text{(A)} = \Big| & \rho^*_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t - \rho^\pi_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
& + \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t \\
& - \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \Big|
\end{aligned}$$

and

$$\begin{aligned}
\text{(B)} = \Big| & \rho^\pi_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t - \rho^\pi_{\omega,n\delta t}\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
& + \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
& - \left[v(d^\pi_{\omega,n\delta t}+1, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \Big|
\end{aligned}$$

We will bound (A), (B), and (C) separately.

- **Bounding (A):** Note that $\text{(A)} = |\text{(A1)} - \text{(A2)}|$ where

$$\begin{aligned}
\text{(A1)} = {} & \rho^*_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t \\
& + \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t \\
\text{(A2)} = {} & \rho^\pi_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
& + \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t
\end{aligned}$$

The only difference between (A1) and (A2) is $\rho^*_{\omega,n\delta t}$ and $\rho^\pi_{\omega,n\delta t}$. Recall from Theorem II.1 that $\rho^*_{\omega,n\delta t} = r^*(d^*_{\omega,n\delta t}, T-n\delta t)$ satisfies the first order condition of the revenue-to-go function for the state $(d^*_{\omega,n\delta t}, T-n\delta t)$. Therefore, following similar steps to when we proved bound (A.45), we can show that (A) is upper bounded by

$$\frac{1}{2}\bar{C}_{xx}\xi(d^*_{\omega,n\delta t})(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t})^2\delta t$$

where

$$\begin{aligned}
\bar{C}_{xx} &:= \sup_r \left| \frac{\partial^2[rx(r)]}{\partial r^2} + \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]x''(r) \right| \\
&\leq C_{xx} + \left| \frac{x(\bar{r})}{x'(\bar{r})} + \bar{r} \right| \cdot \sup_r |x''(r)|,
\end{aligned}$$

where $\bar{r}$ is the price that optimizes the expected revenue-to-go given state $(d^*_{\omega,n\delta t}, T-$

$(n+1)\delta t)$. The inequality follows from $V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t) = -\frac{x(\bar{r})}{x'(\bar{r})} - \bar{r}$ (Theorem II.1) and from Assumption II.1(v). Note that from Assumption II.1(iv), $\frac{x(r)}{x'(r)} + r$ is finite since r is finite.
Hence, (A) is upper bounded by

$$\mathcal{O}\left(\xi(d^*_{\omega,n\delta t})(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t})^2\delta t\right) = \mathcal{O}\left(\xi(d^\pi_{\omega,n\delta t})(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t})^2\delta t\right). \tag{A.55}$$

Here the equality comes from the same argument as when we bounded (A.45) with (A.46).

- **Bounding (B):** Note that (B) $\leq$ (B1) + (B2), where

$$\begin{aligned}
\text{(B1)} &= \left|\rho^\pi_{\omega,n\delta t} x(\rho^\pi_{\omega,n\delta t}) \cdot (\xi(d^*_{\omega,n\delta t}) - \xi(d^\pi_{\omega,n\delta t}))\delta t\right|, \\
\text{(B2)} &= \Big|\left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t}) \\
&\quad - \left[v(d^\pi_{\omega,n\delta t}+1, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^\pi_{\omega,n\delta t})\Big| \cdot x(\rho^\pi_{\omega,n\delta t})\delta t
\end{aligned}$$

We examine the bounds of (B1) and (B2) separately.
Using the same arguments in deriving the bounds (A.47) and (A.49), we have

$$\text{(B1)} \leq C_x \left|\xi(d^\pi_{\omega,n\delta t})\left(1 + \frac{T-\delta t}{e^{T-\delta t}} o(1)\right)\frac{n\delta t}{e^{n\delta t}}\mathcal{O}(1)\,\delta t\right|. \tag{A.56}$$

Furthermore, the order of (B2) cannot exceed

$$\begin{aligned}
&= \mathcal{O}\Big(\big|[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t}) \\
&\quad -[v(d^\pi_{\omega,n\delta t}+1, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)]\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\big|\,\delta t\Big) \qquad \text{(A.57)} \\
&= \mathcal{O}\left(\left|V(d^*_{\omega,n\delta t}, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)\right|\right). \qquad \text{(A.58)}
\end{aligned}$$

Here, (A.57) can be interpreted as the difference between the expected revenue loss of having one more adoption under the optimal policy π^* and that under the policy π. Hence, (A.57) is upper bounded by (A.58), which is the difference between the total expected revenue-to-go under π^* and under π.

From (A.54), and from the bounds (A.55), (A.56) and (A.58), the following constraint must hold almost surely:

$$|V(D^*_{n\delta t}, T-(n+1)\delta t) - v(D^\pi_{n\delta t}, T-(n+1)\delta t)|$$

$$\leq \mathcal{O}\left(\xi(D^\pi_{n\delta t})(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t})^2\delta t\right) + \mathcal{O}\left(\xi(D^\pi_{n\delta t})\frac{n\delta t}{e^{n\delta t}}\delta t\right)$$
$$+ \mathcal{O}\left(|V(D^*_{n\delta t}, T-(n+1)\delta t) - v(D^\pi_{n\delta t}, T-(n+1)\delta t)|\right).$$

From the condition of the proposition, $\mathbb{E}[(r^*_{n\delta t} - r^\pi_{n\delta t})^2] = \Omega\left(\frac{n\delta t}{e^{n\delta t}}\right)$. Therefore, taking the conditional expectation of the above bound given $\mathcal{F}_{n\delta t}$

$$\begin{aligned}
\Psi_n &\leq \mathcal{O}\left(\mathbb{E}\left[\xi(D^\pi_{n\delta t}) \mid \mathcal{F}_{n\delta t}\right]\left(r^*_{n\delta t} - r^\pi_{n\delta t}\right)^2\delta t\right) + \mathbb{E}\left[\Psi_{n+1} \mid \mathcal{F}_{n\delta t}\right] \\
&\leq \mathcal{O}\left(\frac{D^\pi_{n\delta t}+1}{n\delta t + t_0}\left(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t}\right)^2\delta t\right) + \mathbb{E}\left[\Psi_{n+1} \mid \mathcal{F}_{n\delta t}\right] \\
&= \mathcal{O}\left(\mathbb{E}\left[\sum_{s=n}^{N-1}\frac{D^\pi_{s\delta t}+1}{s\delta t + t_0}\left(\rho^\pi_{\omega,s\delta t} - \rho^*_{\omega,s\delta t}\right)^2 \mid \mathcal{F}_{n\delta t}\right]\right)
\end{aligned}$$

Here, the second inequality is due to Claim A.1. The last step is due to the inductive assumption. This finishes the induction proof.

□

A.2.11 Claim A.2 and proof

The claim below is useful for proving Theorem II.2 and Theorem II.3.

Claim A.2. Under Assumption II.2, the upper bound in Proposition II.3 is tight. Specifically,

$$R^* - R(\pi) = \Omega\left(\mathbb{E}\left[\int_0^T \xi(D^*_t)(r^\pi_t - r^*_t)^2 \mathrm{d}t\right]\right).$$

Proof. **Proof of Claim A.2** Using similar logic as the proof of Proposition II.3, we use induction to prove the more general result on a discretized time horizon:

$$\Psi_n = \Omega\left(\mathbb{E}\left[\sum_{s=n}^{N-1}\xi(D^*_{s\delta t})(r^\pi_{s\delta t} - r^*_{s\delta t})^2\delta t \mid \mathcal{F}_{n\delta t}\right]\right). \quad \text{(A.59)}$$

for all $n = 0, 1, \ldots, N-1$. Note that $\Psi_0 = R^* - R(\pi)$. We need to revise the proof of Proposition II.3 in several steps to show (A.59). Following the logic of the proof of Proposition II.3, we first consider the base step $n = N-1$ with δt time remaining. Recall from (A.42) that

$$
\begin{aligned}
&|V(d^*_{\omega,T-\delta t},\delta t) - v(d^\pi_{\omega,T-\delta t},\delta t)| \\
&= \left|\rho^*_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^*_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^\pi_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t\right| \\
&\quad\le \underbrace{\left|\rho^*_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^*_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t\right|}_{\text{(A)}} \\
&\quad + \underbrace{\left|\rho^\pi_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^\pi_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t\right|}_{\text{(B)}}.
\end{aligned}
$$

Let us denote (A') and (B') as the terms inside the absolute values in (A) and (B), respectively. Therefore, we have

$$
|V(d^*_{\omega,T-\delta t},\delta t) - v(d^\pi_{\omega,T-\delta t},\delta t)| = |\text{(A')} + \text{(B')}| \ge |\text{(A')}| - |\text{(B')}| = \text{(A)} - \text{(B)},
$$

where the inequality is due to the triangle inequality: $|x+y| \ge |x| - |y|$. Note that

$$
\begin{aligned}
\text{(A)} &= \left| \int_{\rho^\pi_{\omega,T-\delta t}}^{\rho^*_{\omega,T-\delta t}} \int_{\rho^*_{\omega,T-\delta t}}^{r} \frac{\partial^2[z\xi(d^*_{\omega,T-\delta t})x(z)\delta t]}{\partial z^2} \mathrm{d}z\mathrm{d}r \right| \\
&\ge \frac{1}{2}\underline{C}\xi(d^*_{\omega,T-\delta t})(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t})^2\delta t = \Theta\left(\xi(d^*_{\omega,T-\delta t})(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t})^2\delta t\right).
\end{aligned}
\tag{A.60}
$$

The equality follows from the arguments in (A.45), and the inequality follows from Assumption II.2. Also, from (A.49), we know that

$$
\text{(B)} \le C_x \left| \left(1 + \frac{1}{(T-\delta t)^3} o(1)\right) \frac{T-\delta t}{e^{T-\delta t}} \mathcal{O}\left(\xi(d^\pi_{\omega,T-\delta t})\right)\delta t \right|,
$$

which diminishes fast.

Therefore, combining the arguments above, we have

$$
\begin{aligned}
\Psi_{N-1} &= \mathbb{E}\left[\left|V(D^*_{T-\delta t},\delta t) - v(D^\pi_{T-\delta t},\delta t)\right|\right] \ge \mathbb{E}\left[\text{(A)}\right] - \mathbb{E}\left[\text{(B)}\right] \\
&\ge \Omega\left(\mathbb{E}\left[\xi(D^*_{T-\delta t})(r^\pi_{T-\delta t} - r^*_{T-\delta t})^2\delta t\right]\right) - \frac{T-\delta t}{e^{T-\delta t}} \cdot \mathcal{O}(\mathbb{E}\left[\xi(D^\pi_{T-\delta t})\right])\delta t \\
&= \Omega\left(\mathbb{E}\left[\xi(D^*_{T-\delta t})(r^\pi_{T-\delta t} - r^*_{T-\delta t})^2\delta t\right]\right).
\end{aligned}
$$

The last relationship is due to the condition of Proposition II.3, $\mathbb{E}[(r^\pi_{T-\delta t} - r^*_{T-\delta t})^2] \ge \alpha(\frac{T-\delta t}{e^{T-\delta t}})$. This finishes the base step.

For the inductive step, we assume that the result holds for $n+1$. Specifically,

$$\Psi_{n+1} = \Omega\left(\mathbb{E}\left[\sum_{s=n+1}^{N-1} \xi(D^*_{s\delta t})(r^\pi_{s\delta t} - r^*_{s\delta t})^2 \delta t \mid \mathcal{F}_{(n+1)\delta t}\right]\right). \tag{A.61}$$

We need to show that this implies the result holding for n.

We revise the proof of Proposition II.3 as follows. First, recall from (A.54), we have

$$\begin{aligned}
&|V\left(d^*_{\omega,n\delta t}, T - n\delta t\right) - v\left(d^\pi_{\omega,n\delta t}, T - n\delta t\right)| \\
&\quad = \Big| \rho^*_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t - \rho^\pi_{\omega,n\delta t}\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
&\qquad + \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t \\
&\qquad - \left[v(d^\pi_{\omega,n\delta t}+1, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
&\qquad + V(d^*_{\omega,n\delta t}, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t) \Big|.
\end{aligned}$$

We let (C”) $:= V(d^*_{\omega,n\delta t}, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)$,

$$\begin{aligned}
\text{(A”)} := \; &\rho^*_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t - \rho^\pi_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
&+ \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t \\
&- \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t
\end{aligned}$$

and

$$\begin{aligned}
\text{(B”)} := \; &\rho^\pi_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t - \rho^\pi_{\omega,n\delta t}\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
&+ \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\
&- \left[v(d^\pi_{\omega,n\delta t}+1, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t.
\end{aligned}$$

Then, because of the triangle inequality $|x+y| \geq |x| - |y|$, we know

$$|V(d^*_{\omega,n\delta t}, T-n\delta t) - v(d^\pi_{\omega,n\delta t}, T-n\delta t)| = |\text{(C”)} + \text{(A”)} + \text{(B”)}| \geq |\text{(C”)} + \text{(A”)}| - |\text{(B”)}|. \tag{A.62}$$

Note that (A”) ≥ 0 and (C”) ≥ 0 because (A”) is the difference between the expected revenue during δt under optimal price $\rho^*_{\omega,n\delta t}$

$$\begin{aligned}
&\rho^*_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t \\
&+ \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t
\end{aligned}$$

and the expected revenue under the suboptimal price $\rho^{\pi}_{\omega,n\delta t}$

$$\rho^{\pi}_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^{\pi}_{\omega,n\delta t})\delta t$$
$$+\left[V(d^*_{\omega,n\delta t}+1,T-(n+1)\delta t)-V(d^*_{\omega,n\delta t},T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^{\pi}_{\omega,n\delta t})\delta t,$$

and (C") is the optimal expected revenue minus the expected revenue given the suboptimal price path. Hence, we know the right hand side of (A.62) equals to (C") + (A") − |(B")|.

We know the following holds from the definition of π^*:

$$|(\text{A"})| \tag{A.63}$$
$$=\xi(d^*_{\omega,n\delta t})\delta t \tag{A.64}$$
$$\cdot\left|\int_{\rho^{\pi}_{\omega,n\delta t}}^{\rho^*_{\omega,n\delta t}}\int_{\rho^*_{\omega,n\delta t}}^{r}\left(\frac{\partial^2 zx(z)}{\partial z^2}+[V(d^*_{\omega,n\delta t}+1,T-(n+1)\delta t)-V(d^*_{\omega,n\delta t},T-(n+1)\delta t)]x''(z)\right)\right.$$
$$\left.\mathrm{d}z\mathrm{d}r\right|$$
$$\geq \xi(d^*_{\omega,n\delta t})\delta t\cdot\inf_z\left|\frac{\partial^2}{\partial z^2}(zx(z))-\left(\frac{x(\bar{r})}{x'(\bar{r})}+\bar{r}\right)x''(z)\right|\cdot\int_{\rho^{\pi}_{\omega,n\delta t}}^{\rho^*_{\omega,n\delta t}}\int_{\rho^*_{\omega,n\delta t}}^{r}\mathrm{d}z\mathrm{d}r,$$

where $\bar{r}$ is the price that optimizes the expected revenue-to-go given state $(d^*_{\omega,n\delta t},T-(n+1)\delta t)$. Hence, according to Assumption II.1 and Assumption II.2, we know

$$|(\text{A"})|=\Theta\left(\xi(d^*_{\omega,n\delta t})(\rho^*_{\omega,n\delta t}-\rho^{\pi}_{\omega,n\delta t})^2\delta t\right) \tag{A.65}$$

Note that |(B")| does not exceed the order of |(C")| according to (A.58). Therefore, together with (A.61) and (A.65), we know from taking the conditional expectation of (A.62) given $\mathcal{F}_{n\delta t}$ that:

$$\Psi_n=\Omega\left(\mathbb{E}\left[|V(D^*_{n\delta t},T-(n+1)\delta t)-v(D^{\pi}_{n\delta t},T-(n+1)\delta t)|+\xi(D^*_{n\delta t})\left(r^*_{n\delta t}-r^{\pi}_{n\delta t}\right)^2\delta t\mid\mathcal{F}_{n\delta t}\right]\right)$$
$$=\Omega\left(\mathbb{E}\left[\sum_{s=n}^{N-1}\xi(D^*_{s\delta t})\left(r^{\pi}_{s\delta t}-r^*_{s\delta t}\right)\delta t\mid\mathcal{F}_{n\delta t}\right]\right),$$

where the last equation follows from the inductive hypothesis. **end proof of Claim A.2**

□

A.2.12 Proof of Lemma II.4

Proof. Consider any $t \in (0, T]$. Recall that $r_t^*(\theta, d)$ denotes the Markovian Bass price if t is the elapsed time, d is the cumulative adoptions, and θ is the parameter set. Note that $\theta_0 = (p_0, q_0, m_0)$ is the true parameter set, and $\theta_t = (p_t, q_t, m_t)$ is the parameter set used in policy π as an input to $r_t^*(\cdot, \cdot)$ to determine the price at time t. The following holds almost surely:

$$(r_t^*(\theta_t, D_t^\pi) - r_t^*(\theta_0, D_t^*))^2 = \Theta\left(\underbrace{[r_t^*(\theta_t, D_t^\pi) - r_t^*(\theta_0, D_t^\pi)]^2}_{\text{(A)}}\right) + \Theta\left(\underbrace{[r_t^*(\theta_0, D_t^\pi) - r_t^*(\theta_0, D_t^*)]^2}_{\text{(B)}}\right).$$

Note that Θ is the limiting effect on (A) and (B) as m_0 grows.

Bounding (A): The difference of the two prices in the (A) is due to the parameter difference. We first examine (A). Define $\bar{p} := \max\{p_0, p_t\}$ and $\underline{p} := \min\{p_0, p_t\}$. Since p_0, p_t are positive and finite values, then so are $\bar{p}, \underline{p}$. We similarly define $\bar{q}, \underline{q}, \bar{m}, \underline{m}$. Let $\mathcal{P} := [\underline{p}, \bar{p}] \times [\underline{q}, \bar{q}] \times [\underline{m}, \bar{m}]$.

From the property that $|x - y|^\top (\inf_z \nabla h(z)) \leq h(x) - h(y) \leq |x - y|^\top (\sup_z \nabla h(z))$, we have that

$$\begin{aligned}&\text{(A)}\\ &\leq \Theta\left[\left(\sup_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r_t^*}{\partial p}\right|\right)^2 (p_0 - p_t)^2 + \left(\sup_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r_t^*}{\partial q}\right|\right)^2 (q_0 - q_t)^2 + \left(\sup_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r_t^*}{\partial m}\right|\right)^2 (m_0 - m_t)^2\right],\end{aligned}$$

and

$$\begin{aligned}&\text{(A)}\\ &\geq \Theta\left[\left(\inf_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r_t^*}{\partial p}\right|\right)^2 (p_0 - p_t)^2 + \left(\inf_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r_t^*}{\partial q}\right|\right)^2 (q_0 - q_t)^2 + \left(\inf_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r_t^*}{\partial m}\right|\right)^2 (m_0 - m_t)^2\right],\end{aligned}$$

Here we treat m as a continuous variable.

We first analyze $|\partial r_t^* / \partial p|$. Using the equation (II.3.2) satisfied by $r_t^*(\theta, d)$, we differentiate r_t^* with respect to p and rearranging terms, we get that for any d,

$$\left|\frac{\partial r_t^*(\theta, d)}{\partial p}\right| = \left|\frac{\partial}{\partial p}[V(d, T-t) - V(d+1, T-t)] \Big/ \left(\frac{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}{x'(r_t^*)^2}\right)\right| \quad \text{(A.66)}$$

Note that if we rearrange (A.23), where $g(d,t) = \frac{\partial V(d,t)}{\partial p}$, we have

$$\underbrace{\frac{\partial}{\partial p}[V(d,T-t) - V(d+1,T-t)]}_{\text{(A1)}} = \underbrace{\frac{\partial^2 V(d,T-t)}{\partial p \partial t} \cdot \frac{1}{(m-d)\left(p+\frac{d}{m}q\right)x(r_t^*)}}_{\text{(A2)}} - \frac{x(r_t^*)}{x'(r_t^*)}\frac{1}{p+\frac{d}{m}q}. \tag{A.67}$$

We now examine (A2) on the right-hand side of (A.67). From the HJB equation (A.15), note that $\frac{\partial}{\partial t}V(d,T-t) = J(r_t^*(\theta,d),d,T-t)$ where $J(r,d,t) := r\lambda(d,r) + [V(d+1,t) - V(d,t)]\lambda(d,r)$. Hence, using chain rule, we know

$$\frac{\partial^2 V(d,T-t)}{\partial p \partial t} = -\frac{\partial J(r,d,T-t)}{\partial r}\Big|_{r=r_t^*}\frac{\partial r_t^*}{\partial p} - \frac{\partial J(r,d,T-t)}{\partial p}\Big|_{r=r_t^*} = 0 - \frac{\partial J(r,d,T-t)}{\partial p}\Big|_{r=r_t^*} \leq 0.$$

Moreover, because the partial effect of p on the expected revenue rate cannot exceed the rate when all the remaining population $(m-d)$ directly adopt the product without being affected by the current price r_t^*, we know

$$\frac{\partial^2 V(d,T-t)}{\partial p \partial t} \geq -(m-d)r_t^* x(r_t^*).$$

Hence, from these lower and upper bounds that we derived for $\frac{\partial^2 V(d,T-t)}{\partial p \partial t}$, (A.67) implies that

$$-\frac{r_t^*}{p+\frac{d}{m}q} - \frac{x(r_t^*)}{x'(r_t^*)}\frac{1}{p+\frac{d}{m}q} \leq \text{(A1)} \leq -\frac{x(r_t^*)}{x'(r_t^*)}\frac{1}{p+\frac{d}{m}q}. \tag{A.68}$$

Therefore, we substitute (A.68) into (A.66) to get

$$\left|\frac{\partial r_t^*}{\partial p}\right| \leq \left|\frac{x'(r_t^*)^2}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right| \cdot \left(\left|\frac{x(r_t^*)}{x'(r_t^*)}\frac{1}{p+\frac{d}{m}q}\right| + \left|\frac{r_t^*}{p+\frac{d}{m}q}\right|\right) \leq \frac{M\bar{x}^u}{C_d(\underline{p}+\underline{q})}\Theta(1),$$

where the last inequality follows from Assumption II.1(i), (iv) and since r_t^* does not scale up with the market size m. The latter is because when m_0 grows, the demand process converges to the deterministic Bass model (Proposition II.1). Hence, the optimal price r_t^* should also converge to the optimal price under the deterministic Bass model, which is not affected by the market size.

Using similar arguments as above, we have

$$\left|\frac{\partial r_t^*}{\partial q}\right| \leq \left|\frac{x'(r_t^*)^2}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right| \cdot \left(\left|\frac{x(r_t^*)}{x'(r_t^*)}\frac{\frac{d}{m}}{p+\frac{d}{m}q}\right| + \left|\frac{r_t^*}{p+\frac{d}{m}q}\right|\right) \leq \Theta\left(\frac{M\bar{x}^u}{C_d(\underline{p}+\underline{q})}\right),$$

and

$$\left|\frac{\partial r_t^*}{\partial m}\right| \le \left|\frac{x'(r_t^*)^2}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right| \cdot \left(\left|\frac{x(r_t^*)}{x'(r_t^*)}\frac{p+\frac{d^2}{m^2}q}{(m-d)(p+\frac{d}{m}q)}\right| + \left|\frac{r_t^*}{p+\frac{d}{m}q}\right|\right) \le \frac{M\bar{x}^u}{C_d}\frac{p+q}{p\underline{m}}\Theta(1).$$

Hence, it follows that (A) $= \Theta(\|\theta_t - \theta_0\|^2)$.

Bounding (B): Note that the difference in the two prices in (B) is due to the difference in past sales. Specifically, the model parameters are the same. For any d and $\theta_0 = (p_0, q_0, m_0)$, we define $f_d := (1 - \frac{d}{m_0})\left(p_0 + q_0\frac{d}{m_0}\right)$. From chain rule, we have

$$\begin{aligned}&\mathbb{E}_{\theta_0}\left[|r_t^*(\theta_0, D_t^*) - r_t^*(\theta_0, D_t^\pi)| \mid \mathcal{F}_t\right] \\ &\le \mathbb{E}_{\theta_0}\left[\sup_{d\in[D_t^*\wedge D_t^\pi, D_t^*\vee D_t^\pi]}\left|\frac{\partial r_t^*(\theta_0,d)}{\partial f_d}\right| \cdot \sup_{d\in[D_t^*\wedge D_t^\pi, D_t^*\vee D_t^\pi]}\left|\frac{\partial f_d}{\partial(d/m_0)}\right| \cdot \left|\frac{D_t^\pi}{m_0} - \frac{D_t^*}{m_0}\right| \mid \mathcal{F}_t\right].\end{aligned} \tag{A.69}$$

Note that $\sup_{d\in[0,m_0]}|\partial f_d/\partial(d/m_0)| = (p_0 + q_0)$. Hence, to bound (A.69), we need to evaluate the bound of $|\partial r_t^*/\partial f_d|$.

From (II.2.2), $F_t^r < 1$ for all $t \le T$ and any deterministic price sequence r. Hence, there exists $\delta > 0$ such that $1 - F_t^r > \delta$ for all $t \le T$. One example of δ is $1 - F_T^r > 1 - (p_0+q_0)\bar{x}^u T$ if $(p_0+q_0)\bar{x}^u T < 1$. From (A.14), for any pricing sample path r_ω of policy π^*, $\mathbb{E}(D_t^{r_\omega}/m_0) < F_t^{r_\omega} < 1 - \delta$. Therefore, $\mathbb{E}(1 - D_t^*/m_0) > \delta$, implying that

$$\mathbb{E}_{\theta_0}\left[f_{D_t^*}\right] = \mathbb{E}_{\theta_0}\left[\left(1 - \frac{D_t^*}{m_0}\right)\left(p_0 + q_0\frac{D_t^*}{m_0}\right)\right] > \gamma^\delta := \delta \times p_0.$$

Also since $\mathbb{E}[D_t^*/m_0 \mid \mathcal{F}_t] = \mathbb{E}\left[\int_0^t (1 - \frac{D_s^*}{m_0})(p_0 + q_0\frac{D_s^*}{m_0})x(r_s^*)\mathrm{d}s \mid \mathcal{F}_t\right]$ for all t and the integrand is positive, we have $\mathbb{E}\left(1 - D_t^*/m_0 \mid \mathcal{F}_t\right) > \delta$ as well, so

$$\mathbb{E}_{\theta_0}\left[f_{D_t^*} \mid \mathcal{F}_t\right] = \mathbb{E}_{\theta_0}\left[\left(1 - \frac{D_t^*}{m_0}\right)\left(p_0 + q_0\frac{D_t^*}{m_0}\right) \mid \mathcal{F}_t\right] > \gamma^\delta. \tag{A.70}$$

For any (d,t), we differentiate (II.3.3) by f_d for both sides, which yields

$$\frac{\partial^2 V(d, T-t)}{\partial f_d \partial t}\Big/ m_0 + \frac{x(r_t^*)^2}{x'(r_t^*)} + f_d\frac{2x(r_t^*)x'(r_t^*)^2 - x(r_t^*)^2x''(r_t^*)}{x'(r_t^*)^2}\frac{\partial r_t^*}{\partial f_d} = 0. \tag{A.71}$$

Then,

$$\gamma^\delta \cdot \inf_r \left|\frac{2x(r)x'(r)^2 - x(r)^2x''}{x'(r)^2}\right| \cdot \mathbb{E}_{\theta_0}\left(\left|\frac{\partial r_t^*(\theta_0, D_t^*)}{\partial f_d}\right| \mid \mathcal{F}_t\right) \tag{A.72}$$

$$\leq \mathbb{E}_{\theta_0}\left(f_{D_t^*}\cdot\left|\frac{2x(r_t^*)x'(r_t^*)^2-x(r_t^*)^2x''(r_t^*)}{x'(r_t^*)^2}\cdot\frac{\partial r_t^*(\theta_0,D_t^*)}{\partial f_d}\right| \mid \mathcal{F}_t\right) \tag{A.73}$$

$$= \mathbb{E}_{\theta_0}\left(\left|\frac{\partial^2 V(D_t^*,T-t)}{\partial f_d\partial t}/m_0+\frac{x(r_t^*(\theta_0,D_t^*))^2}{x'(r_t^*(\theta_0,D_t^*))}\right| \mid \mathcal{F}_t\right) \tag{A.74}$$

$$= \mathbb{E}_{\theta}\left(\left|\frac{\partial^2 V(D_t^*,T-t)}{\partial f_d\partial t}/m_0+\frac{\partial V(D_t^*,T-t)}{\partial t}/(f_{D_t^*}m_0)\right| \mid \mathcal{F}_t\right) \tag{A.75}$$

where (A.73) follows from (A.70), (A.74) follows from (A.71), (A.75) is from (II.3.3).

Note that $V(D_t^*,T-t)=\mathbb{E}_{\theta_0}\left(\int_0^t m_0 f_{D_s^*}x(r_s^*(\theta_0,D_s^*))r_s^*(\theta_0,D_s^*)\mathrm{d}s \mid \mathcal{F}_t\right)$. Hence,

$$\mathbb{E}_{\theta_0}\left(\frac{\partial^2 V(D_t^*,T-t)}{\partial f_d\partial t}\Big/m_0 \mid \mathcal{F}_t\right)=r_t^*(\theta_0,D_t^*)x(r_t^*(\theta_0,D_t^*)),$$
$$\mathbb{E}_{\theta_0}\left(\frac{\partial V(D_t^*,T-t)}{\partial t}\Big/(f_{D_t^*}m_0) \mid \mathcal{F}_t\right)=r_t^*(\theta_0,D_t^*)x(r_t^*(\theta_0,D_t^*)),$$

so an upper bound for (A.72) is $2\sup_r|rx(r)|\leq 2C_x$ from Assumption II.1(v). Moreover, also from Assumption II.1, we can show that $\inf_r\left|\frac{2x(r)x'(r)^2-x(r)^2x''(r)}{x'(r)^2}\right|\geq\frac{C_d\bar{x}^l}{M^2}$, then

$$\mathbb{E}_{\theta_0}\left(\left|\frac{\partial r_t^*(\theta_0,D_t^*)}{\partial f_d}\right| \mid \mathcal{F}_t\right)\leq\frac{2C_xM^2}{\gamma^\delta C_d\bar{x}^l}. \tag{A.76}$$

Using the same arguments, we can also get the same upper bound for $d=D_t^\pi$, and also for any $d\in[D_t^\pi\wedge D_t^*,\ D_t^\pi\vee D_t^*]$.

Hence, from (A.69) and (A.76), we have

$$\mathbb{E}_{\theta_0}\left[|r_t^*(\theta_0,D_t^*)-r_t^*(\theta_0,D_t^\pi)| \mid \mathcal{F}_t\right]\leq\frac{2C_xM^2}{\gamma^\delta C_d\bar{x}^l}\cdot(p_0+q_0)\cdot\mathbb{E}_{\theta_0}\left[\left|\frac{D_t^\pi}{m_0}-\frac{D_t^*}{m_0}\right| \mid \mathcal{F}_t\right]$$

Therefore, it follows that

$$\begin{aligned}&\mathbb{E}\left[(r_t^*(\theta_0,D_t^*)-r_t^*(\theta_0,D_t^\pi))^2 \mid \mathcal{F}_t\right]=\mathcal{O}\left(\mathbb{E}_{\theta_0}\left[\left(\frac{D_t^*}{m_0}-\frac{D_t^\pi}{m_0}\right)^2 \mid \mathcal{F}_t\right]\right)\\&\leq\alpha\left(\frac{t}{e^t}\right)^2\mathbb{E}\left[(r_t^*(\theta_0,D_t^*)-r_t^*(\theta_0,D_t^\pi))^2 \mid \mathcal{F}_t\right]+\mathcal{O}\left(\frac{1}{m_0}\right)\end{aligned} \tag{A.77}$$

for some $\alpha>0$ independent of m_0, where the last relationship follows due to Lemma A.2. This concludes our bound on (B).

Therefore, from the bounds on (A) and (B), we can conclude that

$$\mathbb{E}\left[(r_t^*(\theta_0, D_t^*) - r_t^*(\theta_t, D_t^\pi))^2 \mid \mathcal{F}_t\right] = \Theta\left(\mathbb{E}\left[\|\theta_t - \theta_0\|^2 \mid \mathcal{F}_t\right]\right) + \mathcal{O}\left(\frac{1}{m_0}\right).$$

□

A.2.13 Proof of Theorem II.2

The Bayesian Cramer-Rao bound will be useful in our proof of Theorem II.2. It states that, under some regularity conditions, the distribution of an estimator of an absolutely continuous function g of θ cannot have a variance less than the classical informational bound.

Lemma A.4 (Bayesian Cramer-Rao bound.)**.** Let $\{f(\cdot \mid \theta) : \theta \in \Theta\}$ be a family of probability density functions on some sample space $\mathcal{X}$, where the parameter space Θ is a closed interval on the real line. Let $\mu(\theta)$ be some probability density on $\theta \in \Theta$. Suppose that μ and $f(x \mid \cdot)$ are both absolutely continuous, and that μ converges to zero at the endpoints of the interval Θ. If X is the random sample, let $\hat{g}(X)$ denote an estimator of $g(\theta)$, where $g : \Theta \mapsto \mathbb{R}$ is an absolutely continuous function. Then,

$$\mathbb{E}_\theta\left[(\hat{g}(X) - g(\theta))^2\right] \geq \frac{\left(\mathbb{E}_\theta\left[\frac{\mathrm{d}}{\mathrm{d}\theta} g(\theta)\right]\right)^2}{\mathbb{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta} \ln f(X \mid \theta)\right)^2\right] + \mathbb{E}\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta} \ln \mu(\theta)\right)^2\right]}$$

where $\mathbb{E}_\theta[\cdot]$ denotes the expectation with respect to the joint distribution of $f(X \mid \theta)$ and $\mu(\theta)$.

Proof. Proof of Theorem II.2.

To prove the lower bound, we only need to consider the case where one parameter, say q_0, is unknown. This is because more unknown and independent parameters can only worsen the revenue loss. Hence, we assume only $\theta_0 = q_0$ is unknown.

First, using the Bayesian Cramer-Rao inequality (Lemma A.4), we show the following claim which is a lower bound on the pricing error for any pricing-and-learning policy $\tilde{\pi} \in \Pi$. (In this proof, we use the tilde notation to distinguish the policy $\tilde{\pi}$ from the mathematical constant π.)

Claim A.3. Suppose $x(r) = e^{-r}$ for $r \in [0, 2)$. Let $\theta = q_0$ be a random variable taking values in $\Theta = \left[\frac{1}{4}, \frac{5}{4}\right]$ with the density $\mu(\theta) = 2[\cos(\pi(\theta - 3/4))]^2$. Then for any pricing-and-learning policy $\tilde{\pi} \in \Pi$,

$$\mathbb{E}_\theta \left[(r_t^{\tilde{\pi}} - r_t^*)^2 \mid \mathcal{F}_t \right] \geq \alpha \left(\frac{1}{D_t^{\tilde{\pi}}} \right) \tag{A.78}$$

for some $\alpha > 0$ independent of m_0.

Proof. Proof of Claim A.3. For some $t \in (0, T)$, let X denote the sample path at time t under policy $\tilde{\pi}$. Specifically, $X = (D_s, s \in [0, t])$, where we drop the superscript $\tilde{\pi}$ to simplify notation. Using the notation of Lemma A.4, the density function given $\mathcal{F}_t$ is

$$f(X \mid \theta) = \prod_{i=0}^{D_t} f_i(\theta),$$

where $f_i(\theta)$, $i = 0, 1, \ldots, D_t$ are defined in (II.4.2). With abuse of notation, we also set $g(\theta) = r_t^*(\theta)$ and $\hat{g}(X) = r_t^{\tilde{\pi}}(X)$.

We will first bound $\mathbb{E}_\theta \left[\frac{\mathrm{d}}{\mathrm{d}\theta} g(\theta) \right]$. Since $x(r) = e^{-r}$ for $r \in [0, 2)$, then according to Theorem II.1, we have that

$$r_t^*(\theta) = 1 + [V(D_t^*, T - t; \theta) - V(D_t^* + 1, T - t; \theta)],$$

where $V(d, t; \theta)$ has the closed-form expression given in (II.1). Therefore, if $D_t^* = d$, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\theta} r_t^*(\theta) &= \frac{\partial \left[V(d, T - t) - V(d + 1, T - t) \right]}{\partial \theta} \\
&\geq \alpha_1 \left(\frac{\partial}{\partial q_0} \ln \left(1 + \frac{\frac{\Pi_{i=d}^{m_0 - 1} (m_0 - i) {q_0}^{m_0 - d} t^{m_0 - d}}{(m_0 - d)! e^{m_0 - d}}}{1 + \sum_{j=1}^{m_0 - d - 1} \Pi_{i=d}^{d + j - 1} (m_0 - i) \frac{{q_0}^j t^j}{j! e^j}} \right) \right) \\
&= \alpha_2 \left((m_0 - d) \ln q_0 \right) \geq \alpha_2 \ln q_0
\end{aligned}$$

for some $\alpha_1 > 0, \alpha_2 > 0$ independent of m_0. Therefore, we know that

$$\left(\mathbb{E}_\theta \left[\frac{\mathrm{d}}{\mathrm{d}\theta} g(\theta) \right] \right)^2 = \left(\mathbb{E}_\theta \left[\frac{\mathrm{d}}{\mathrm{d}\theta} r_t^*(\theta) \right] \right)^2 \geq \Omega(1). \tag{A.79}$$

Also, we have

$$\mathbb{E}_\theta \left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta} \ln \mu(\theta) \right)^2 \right] = \mathbb{E}_\theta \left[16\pi^2 \left(\cos(\pi(\theta - 3/4)) \sin(\pi(\theta - 3/4)) \right)^2 \right] \leq 16\pi^2. \tag{A.80}$$

To bound $\mathbb{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln f(X\mid\theta)\right)^2\right]$, we use the following standard result (Cover, 1999):

$$\mathbb{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln f(X\mid\theta)\right)^2\mid\mathcal{F}_t\right]=-\mathbb{E}_\theta\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\ln f(X\mid\theta)\mid\mathcal{F}_t\right]=-\mathbb{E}_\theta\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\sum_{i=0}^{D_t}\ln f_i(\theta)\mid\mathcal{F}_t\right]$$
$$=\sum_{i=0}^{D_t}-\mathbb{E}_\theta\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\ln f_i(\theta)\mid\mathcal{F}_t\right].$$

Thus,

$$\mathbb{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln f(X\mid\theta)\right)^2\mid\mathcal{F}_t\right]=\sum_{i=0}^{D_t}-\mathbb{E}_\theta\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\ln f_i(\theta)\mid\mathcal{F}_t\right]\le\sum_{i=0}^{D_t}\frac{1}{1\cdot {q_0}^2}=(D_t+1)/{q_0}^2. \tag{A.81}$$

Hence, taking (A.79),(A.80),(A.81) into Lemma A.4, we have

$$\mathbb{E}_\theta\left[(r_t^{\tilde{\pi}}-r_t^*)^2\mid\mathcal{F}_t\right]=\mathbb{E}_\theta\left[(\hat{g}(X)-g(\theta))^2\mid\mathcal{F}_t\right]\ge\alpha_3\left(\frac{1}{D_t+1+16\pi^2}\right)=\alpha\left(\frac{1}{D_t+1}\right) \tag{A.82}$$

for some $\alpha>0,\alpha_3>0$ independent of m_0.

end proof of Claim A.3 ■ □

Next, we want to apply Claim A.2 to prove the lower bound on regret. Thus, we need to check whether Assumption II.2 and the condition of Proposition II.3 hold. Notice that, $\left|\frac{\partial^2}{\partial r^2}[re^{-r}]\right|=(2-r)e^{-r}$, so Assumption II.2 holds. Also, from Claim A.3, we know $\mathbb{E}\left[(r_t^{\tilde{\pi}}-r_t^*)^2\mid\mathcal{F}_t\right]\ge\alpha\left(\frac{1}{D_t^{\tilde{\pi}}+1}\right)$. Since

$$\mathbb{E}\left(\frac{1}{D_t^{\tilde{\pi}}+1}\mid\mathcal{F}_t\right)=\frac{1}{\int_0^t\left(m_0-D_s^{\tilde{\pi}}\right)\left(p_0+q_0\frac{D_s^{\tilde{\pi}}}{m_0}\right)\mathrm{d}s+1}\ge\frac{1}{\bar{x}^u m_0(p_0+q_0)t}\ge\alpha' te^{-t}/m_0,$$

for some α' independent of t and m_0. This implies that the condition of Proposition II.3 is satisfied. Hence, from Claim A.2, we have

$$R^*-R(\tilde{\pi})=\Omega\left(\mathbb{E}\left[\int_0^T\xi(D_t^*)(r_t^{\tilde{\pi}}-r_t^*)^2\mathrm{d}t\right]\right)$$
$$=\Omega\left(\mathbb{E}\left[\int_0^T\xi(D_t^{\tilde{\pi}})(r_t^{\tilde{\pi}}-r_t^*)^2\mathrm{d}t\right]\right) \tag{A.83}$$

$$= \Omega\left(\mathbb{E}\left[\int_0^T \xi(D_t^{\tilde{\pi}})\frac{1}{D_t^{\tilde{\pi}}+1}\mathrm{d}t\right]\right) \tag{A.84}$$

where (A.83) comes from the same analysis of (A.46) by replacing $\xi(D_t^*)$ by $\frac{\xi(D_t^*)}{\xi(D_t^{\tilde{\pi}})}\xi(D_t^{\tilde{\pi}})$, and (A.84) comes from Claim A.3.

Then, we finally prove the lower bound on regret. Since $\mathbb{E}(D_t^{\tilde{\pi}}+1 \mid \mathcal{F}_t) \leq \max_{s\leq t}\xi(D_s^{\tilde{\pi}})\cdot 1\cdot(t+t_0)$ with $t_0 = \Theta\left(m_0^{-1}\right)$ (we add t_0 here to avoid meaningless cases where $t=0$), we know

$$\begin{aligned}(\text{A.84}) &\geq \alpha\left(\mathbb{E}\left[\int_0^T \frac{\xi(D_t^{\tilde{\pi}})}{\max_{s\leq t}\xi(D_s^{\tilde{\pi}})}\frac{1}{t+t_0}\mathrm{d}t\right]\right)\\ &\geq \alpha\left(\mathbb{E}\left[\int_0^T \frac{\gamma^\delta}{(p_0+q_0)^2/(4q_0)}\frac{1}{t+t_0}\mathrm{d}t\right]\right) = \Omega(\ln m_0)\end{aligned} \tag{A.85}$$

for some $\alpha > 0$ independent of m_0, where the second inequality comes from (A.70).

This concludes our proof. □

A.2.14 Proof of Theorem II.3

Proof. To simplify notation in this proof, we refer to $\hat{\theta}_0$ as θ instead. Let $D^\pi = (D_t^\pi, t\geq 0)$ denote the cumulative adoption process under policy π that offers the price $r_t^*(\theta,d)$ when the state is (d,t). Recall that $r_t^*(\theta,d)$ is the Markovian Bass price (Theorem II.1) under parameter set θ and state (d,t).

Note that $R^* - R(\pi)$ can be interpreted as the expected revenue loss of mis-specifying the demand parameter as θ, when the true parameter is θ_0. From Lemma II.4, we know that $\mathbb{E}[|r_t^\pi - r_t^*|] = \Omega(te^{-t}{m_0}^{-1})$. Since Assumption II.2 holds, then according to Claim A.2, we have

$$\begin{aligned}R^* - R(\pi) &\geq \mathbb{E}_{\theta_0}\left[\int_0^T \xi(D_t^*)(r_t^\pi - r_t^*)^2\mathrm{d}t\right]\\ &\geq m_0\left[\min_{F\in[0,1]}(1-F)(p_0+q_0F)\right]\cdot\mathbb{E}_{\theta_0}\left[\int_0^T \mathbb{E}_{\theta_0}\left[(r_t^\pi - r_t^*)^2 \mid \mathcal{F}_t\right]\mathrm{d}t\right]\\ &\geq m_0\left[\min_{F\in[0,1]}(1-F)(p_0+q_0F)\right]\cdot\int_0^T \Theta\left(\|\theta-\theta_0\|^2\right)\mathrm{d}t\end{aligned}$$

$$= \mathcal{E}^2 T\Omega(m_0).$$

The second inequality is because by definition $\xi(d) = m_0\left(1 - \frac{d}{m_0}\right)\left(p_0 + q_0\frac{d}{m_0}\right)$ for any d. The third inequality is because of Lemma II.4. The equality is due to $\|\theta_0 - \theta\|^2 = \mathcal{E}^2$. □

A.2.15 Proof of Theorem II.4

Proof. Recall that $\theta_0 = (p_0, q_0, m_0)$ denotes the true parameter vector. For notational convenience, we will use π to denote the MBP-MLE policy π^{M}. Consequently, we will denote the price process and the demand process under MBP-MLE as $r^\pi = (r_t^\pi, t \geq 0)$ and $D^\pi = (D_t^\pi, t \geq 0)$, respectively. The price process under the hindsight optimal policy is $r^* = (r_t^*, t \geq 0)$.

We will use Proposition II.3 and Lemma II.4 to prove the theorem. Therefore, we need to check whether the conditions required in Proposition II.3 and Lemma II.4 are satisfied.

The condition required in Lemma II.4 is that $\hat{p}_t$, $\hat{q}_t$ and $\hat{m}_t$ are finite values, and that $\hat{p}_t + \hat{q}_t > 0$ and $m_t > 0$. This can be observed from checking the likelihood function ℓ_t. Since $\widehat{D}_t \geq 3$, we know that $\hat{m}_t > \widehat{D}_t - 1$, because otherwise, the likelihood function is either 0 or negative. If both $\hat{p}_t$ and $\hat{q}_t$ are zero, the likelihood function is 0. If either $\hat{p}_t = +\infty$ or $\hat{q}_t = +\infty$ or $\hat{m}_t = +\infty$, then the likelihood function is 0. Therefore, the ML estimates satisfy the condition of Lemma II.4.

We check that $\mathbb{E}[|r_t^\pi - r_t^*|] \geq \alpha t e^{-t} {m_0}^{-1}$ for all $t \in [0, T]$, which is the condition needed for Proposition II.3. Conditional on $\mathcal{F}_t$, when the state is (d, t), policy π offers price $r_t^*(\hat{\theta}_t, d)$, which is the Markovian Bass price when the parameter set is the ML estimate $\hat{\theta}_t = \hat{\theta}_t(\widehat{\mathbf{U}}_t)$. Hence,

$$\mathbb{E}[|r_t^\pi - r_t^*|] = \mathbb{E}\left[\mathbb{E}\left[|r_t^\pi - r_t^*| \mid \mathcal{F}_t\right]\right] = \Omega\left(\mathbb{E}\left[\mathbb{E}\left[\left\|\hat{\theta}_t - \theta_0\right\| \mid \mathcal{F}_t\right]\right]\right) = \Omega\left(\mathbb{E}\left(\frac{1}{\sqrt{D_t^\pi + 1}}\right)\right)$$
$$\geq \frac{1}{m_0} \geq \alpha t e^{-t} {m_0}^{-1}.$$

Here, the first bound comes from Lemma II.4. The second bound comes from Lemma II.3. Hence, the condition needed for Proposition II.3 is met.

To prove $R^* - R(\pi) \leq \mathcal{O}(\ln m_0)$, according to Proposition II.3, it suffices to show that

$$\mathbb{E}\left[\int_0^T \frac{D_t^\pi + 1}{t + t_0}(r_t^\pi - r_t^*)^2 \mathrm{d}t\right] = \mathcal{O}(\ln m_0).$$

We know from Lemma II.3 that, for any $t \in (0, T]$, the conditional expected estimation errors are

$$\mathbb{E}\left((\hat{p}_t - p_0)^2 \mid \mathcal{F}_t\right) = \mathcal{O}\left((D_t^\pi + 1)^{-1}\right), \tag{A.86}$$

$$\mathbb{E}\left((\hat{q}_t - q_0)^2 \mid \mathcal{F}_t\right) = \mathcal{O}\left((D_t^\pi + 1)^{-1}\right), \tag{A.87}$$

$$\mathbb{E}\left((\hat{m}_t - m_0)^2 \mid \mathcal{F}_t\right) = \mathcal{O}\left((D_t^\pi + 1)^{-1}\right). \tag{A.88}$$

Also, we know $(D_t^\pi + 1)^{-1}$ dominates $\mathcal{O}(1/m_0)$. This is because

$$\frac{1}{m_0} \leq \mathcal{O}\left((D_t^\pi + 1)^{-1}\right).$$

Then by Lemma II.4, it suffices to show that

$$\mathbb{E}\left[\int_0^T \frac{D_t^\pi + 1}{t + t_0} \|\hat{\theta}_t - \theta_0\|^2 \mathrm{d}t\right] = \mathcal{O}\left(\ln(m_0 T + 1)\right).$$

Notice that the initial revenue loss when $D_t^\pi \leq 3$ is at most $\ln(m_0)$. Then, by conditioning on $\mathcal{F}_t$,

$$\begin{aligned}&\mathbb{E}\left(\mathbb{E}\left[\int_0^T \frac{D_t^\pi + 1}{t + t_0} |\hat{p}_t - p_0|^2 \mathrm{d}t \mid \mathcal{F}_T\right]\right) + \mathbb{E}\left(\mathbb{E}\left[\int_0^T \frac{D_t^\pi + 1}{t + t_0} |\hat{q}_t - q_0|^2 \mathrm{d}t \mid \mathcal{F}_T\right]\right) \\ &+ \mathbb{E}\left(\mathbb{E}\left[\int_0^T \frac{D_t^\pi + 1}{t + t_0} |\hat{m}_t - m_0|^2 \mathrm{d}t \mid \mathcal{F}_T\right]\right) \leq \mathbb{E}\left(\int_0^T \frac{1}{t + t_0} \mathrm{d}t\right) = \mathcal{O}(\ln(m_0)),\end{aligned}$$

where the last inequality follows from (A.86)–(A.88). This proves the theorem. □

A.2.16 Proof of Theorem II.5

Proof. For convenience, we will use π to denote the MBP-MLE-Limited policy $\pi^{\text{M-Lim}}$. Recall that $\hat{\theta}_t = \hat{\theta}_t(\widehat{\mathbf{U}}_t)$ denotes the ML estimator of the parameter set, given data $\widehat{\mathbf{U}}_t$. Note that $\hat{\theta}_t$ influences the policy only if t is a price change epoch. We will denote $\hat{\theta}^\pi = (\hat{\theta}_t^\pi, t \geq 0)$ as the parameter process under MBP-MLE-Limited, where $\hat{\theta}_t^\pi$ is equal to the ML estimator at the most recent price change epoch. Given state (d, t), recall that $r_t^*(\theta_0, d)$ denotes the Markovian Bass price when the demand parameter set is θ_0. We will denote by $r_t^\pi(\hat{\theta}_t^\pi, d)$ the price offered under MBP-MLE-Limited given state (d, t). We will denote the demand process under MBP-MLE-Limited as $D^\pi = (D_t^\pi, t \geq 0)$. The demand process under the hindsight optimal policy is $D^* = (D_t^*, t \geq 0)$.

We will use Proposition II.3 to prove the theorem. Note that the condition of the proposition that $\mathbb{E}[(r_t^\pi - r_t^*)^2] = \Omega(te^{-t})$ for all t is satisfied. This is because the pricing error of MBP-MLE-Limited is larger than that of MBP-MLE, and we showed that the condition is true for the latter policy when proving Theorem II.4 in Section A.2.15.

According to Proposition II.3, we know it suffices to examine the bound for

$$\mathbb{E}\left[\mathbb{E}\left(\int_0^T \frac{D_t^\pi + 1}{t + t_0} (r_t^* - r_t^\pi)^2 \,\mathrm{d}t \mid \mathcal{F}_T\right)\right]. \tag{A.89}$$

With probability 1, we can decompose the pricing error as follows:

$$\begin{aligned}
&\left(r_t^*(\theta_0, D_t^*) - r_t^\pi(\hat{\theta}_t^\pi, D_t^\pi)\right)^2 \\
&= \left(r_t^*(\theta_0, D_t^*) - r_t^*(\theta_0, D_t^\pi) + r_t^*(\theta_0, D_t^\pi) - r_t^\pi(\theta_0, D_t^\pi) + r_t^\pi(\theta_0, D_t^\pi) - r_t^\pi(\hat{\theta}_t^\pi, D_t^\pi)\right)^2 \\
&= \Theta\left(\underbrace{|r_t^*(\theta_0, D_t^*) - r_t^*(\theta_0, D_t^\pi)|^2}_{\text{(A)}}\right) + \Theta\left(\underbrace{|r_t^*(\theta_0, D_t^\pi) - r_t^\pi(\theta_0, D_t^\pi)|^2}_{\text{(B)}}\right) \\
&+ \Theta\left(\underbrace{\left|r_t^\pi(\theta_0, D_t^\pi) - r_t^\pi(\hat{\theta}_t^\pi, D_t^\pi)\right|^2}_{\text{(C)}}\right),
\end{aligned}$$

because of the triangle inequality.

Similar to how we proved Lemma II.4 (Section A.2.12), specifically from (A.77), taking the expectation of (A) conditioning on $\mathcal{F}_t$ is bounded by

$$\frac{1}{m_0} = \mathcal{O}\left(\mathbb{E}\left[\|\hat{\theta}^\pi - \theta_0\|^2 \mid \mathcal{F}_t\right]\right),$$

where the equality is from Lemma II.3. According to Lemma II.4, (C) is also bounded by $\mathcal{O}\left(\mathbb{E}\left[\|\hat{\theta}^\pi - \theta_0\|^2 | \mathcal{F}_t\right]\right)$.

Let us consider the expected cumulative regret (during one price cycle) resulting from (B) when the true parameter set θ_0 is used by the policy π. Specifically, suppose that t is the start of a price cycle whose length is the time until the next c_t adoptions. Specifically, MBP-MLE-Limited sets the price r_t^π for the entire price cycle, which it computes from the deterministic equivalent of the optimal prices $(r_1, r_2, \ldots, r_{c_t})$ and inter-adoption times $(\Delta t_1, \Delta t_2, \ldots, \Delta t_{c_t})$, as described in Section II.4.2.2. If the cycle's inter-adoption times under π and π^* are equal to $(\Delta t_1, \Delta t_2, \ldots, \Delta t_{c_t})$, then the revenue loss only comes from π using a constant price during a price change epoch, instead of using flexible prices by π^*. In this case, $r_{t+\tau_{i-1}}^* = r_i$ where $\tau_{i-1} := \sum_{k=1}^{i-1} \Delta t_k$ is the time elapsed after the $(i-1)$th adoption in the cycle. Hence, the revenue loss due to (B) is zero since

$$\sum_{i=1}^{c_t} r^*_{t+\tau_{i-1}}\xi(D^\pi_t+i-1)x(r^*_{t+\tau_{i-1}})\Delta t_i - \sum_{i=1}^{c_t} r^\pi_t\xi(D^\pi_t+i-1)x(r^\pi_t)\Delta t_i$$
$$=\sum_{i=1}^{c_t} r_i\xi(D^\pi_t+i-1)x(r_i)\Delta t_i - \left(\sum_{i=1}^{c_t}\xi(D^\pi_t+i-1)\Delta t_i\right)\left(\frac{\sum_{j=1}^{c_t} r_j x(r_j)\xi(D^\pi_t+j-1)\Delta t_j}{\sum_{j=1}^{c_t}\xi(D^\pi_t+j-1)\Delta t_j}\right)=0.$$

where the first equality is from (II.4.7). However, revenue loss will not be zero because of the error from approximating the inter-adoption times. We use $\widehat{D}^\pi$ to denote the corresponding demand sequence from the approximated inter-adoption times. Specifically, under process $\widehat{D}^\pi$, we have $\widehat{D}^\pi_t = D^\pi_t$ and an additional adoption after Δt_1, after Δt_2, and so on. Hence, along with the analysis above, (B) is bounded above in the order of $\left(r^*_t(\theta_0, D^\pi_t) - r^*_t(\theta_0, \widehat{D}^\pi_t)\right)^2$, which is in the same order as (A).

Now, taking the bounds on (A), (B), and (C) into (A.89), to prove the theorem, it suffices to bound

$$\mathbb{E}\left[\int_0^T \frac{D^\pi_t+1}{t+t_0}\left[(p_0-\hat{p}^\pi_t)^2+(q_0-\hat{q}^\pi_t)^2+(m-\hat{m}^\pi_t)^2\right]\mathrm{d}t\big|\mathcal{F}_T\right]. \tag{A.90}$$

Define t_i as the earliest time between T and the occurence of the ith adoption under policy π. Recall that C_i is the number of adoptions in price cycle C_i under π, and $C_{[i]} := \sum_{k=0}^{i} C_i$. Furthermore, the ML estimator $\hat{\theta}^\pi_t$ is only updated at the start of each price cycle. Hence, using Lemma II.3, we know that on any demand sample path, (B.1) can be bounded by

$$\int_0^{t_1}\frac{1}{t+t_0}\mathrm{d}t+\int_{t_1}^{t_2}\frac{1+1}{t+t_0}\mathrm{d}t+\int_{t_2}^{t_3}\frac{2+1}{t+t_0}\mathrm{d}t+\cdots+\int_{t_{C_0-1}}^{t_{C_0}}\frac{C_0-1+1}{t+t_0}\mathrm{d}t$$
$$+\int_{t_{C_{[0]}}}^{t_{C_{[0]}+1}}\frac{C_{[0]}+1}{t+t_0}\frac{1}{C_{[0]}+1}\mathrm{d}t+\int_{t_{C_{[0]}+1}}^{t_{C_{[0]}+2}}\frac{C_{[0]}+1+1}{t+t_0}\frac{1}{C_{[0]}+1}\mathrm{d}t+\cdots+\int_{t_{C_{[0]}+C_1-1}}^{t_{C_{[0]}+C_1}}\frac{C_{[0]}+C_1-1+1}{t+t_0}\frac{1}{C_{[0]}+1}\mathrm{d}t$$
$$+\cdots$$
$$+\int_{t_{C_{[K-1]}}}^{t_{C_{[K-1]}+1}}\frac{C_{[K-1]}+1}{t+t_0}\frac{1}{C_{[K-1]}+1}\mathrm{d}t+\cdots+\int_{t_{C_{[K-1]}+C_K-1}}^{t_{C_{[K-1]}+C_K}}\frac{C_{[K-1]}+C_K-1+1}{t+t_0}\frac{1}{C_{[K-1]}+1}\mathrm{d}t$$
$$\le\int_0^{t_{C_{[0]}}}\frac{C_{[0]}}{t+t_0}\mathrm{d}t+\int_{t_{C_{[0]}}}^{t_{C_{[1]}}}\frac{C_{[1]}}{t+t_0}\frac{1}{C_{[0]}+1}\mathrm{d}t+\int_{t_{C_{[1]}}}^{t_{C_{[2]}}}\frac{C_{[2]}}{t+t_0}\frac{1}{C_{[1]}+1}\mathrm{d}t+\ldots+\int_{t_{C_{[K-1]}}}^{t_{C_{[K]}}}\frac{C_{[K]}}{t+t_0}\frac{1}{C_{[K-1]}+1}\mathrm{d}t$$
$$\le\left(\max\left\{C_0, 1+\max_{i=1,2,\ldots,K}\frac{C_i}{C_{i-1}}\right\}\right)\int_0^T\frac{1}{t+t_0}\mathrm{d}t=\mathcal{O}\left(\left(\max\left\{C_0,1+\max_{i=1,2,\ldots,K}\frac{C_i}{C_{i-1}}\right\}\right)\ln(m_0T)\right).$$

Here, the last inequality is because, for any $i = 1, 2, \ldots, K$,

$$\frac{C_{[i]}}{C_{[i-1]} + 1} = \frac{C_0 + C_1 + \ldots + C_{i-1} + C_i}{C_0 + C_1 + \ldots + C_{i-1} + 1} \le 1 + \frac{C_i}{C_0 + \ldots + C_{i-1} + 1} \le 1 + \frac{C_i}{C_{i-1}}$$

Therefore, we can conclude that

$$R^* - R(\pi) = \mathcal{O}\left(\max\left\{C_0,\ 1 + \max_{i=1,2,\ldots,K} \frac{C_i}{C_{i-1}}\right\} \cdot \ln m_0\right).$$

□

A.2.17 Proof of Proposition II.4

Proof. We only need to analyze the concavity of the first term in (II.5.4), since the remaining terms are linear in μ'. We denote the first term as $\phi(\mu') := \sum_{i=0}^{\widehat{D}_t - 1} \ln {\mu'}^\top y^{i,t_{i+1}}$. In what follows, we will show that $\phi(\mu')$ is strictly and jointly concave in μ'.

To show $\phi(\mu')$ is strictly concave in μ', we need to show that its Hessian is negative definite. For any $k = 1, 2, \ldots, 3(n+1)$, $\ell = 1, 2, \ldots, 3(n+1)$, we have

$$\frac{\partial}{\partial \mu'_k} \phi(\mu') = \sum_{i=0}^{\widehat{D}_t - 1} \frac{y_k^{i,t_{i+1}}}{{\mu'}^\top y^{i,t_{i+1}}}, \qquad \frac{\partial^2}{\partial \mu'_k \mu'_\ell} \phi(\mu') = -\sum_{i=0}^{\widehat{D}_t - 1} \frac{y_k^{i,t_{i+1}} y_\ell^{i,t_{i+1}}}{({\mu'}^\top y^{i,t_{i+1}})^2}.$$

Therefore, for any vector $\mathbf{z} \in \mathbb{R}^{3(n+1)}$, we have

$$\mathbf{z}^\top \nabla^2_{\mu'} \phi(\mu') \mathbf{z} = -\sum_{i=0}^{\widehat{D}_t - 1} \frac{(\mathbf{z}^\top y^{i,t_{i+1}})^2}{({\mu'}^\top y^{i,t_{i+1}})^2} \le 0. \tag{A.91}$$

Hence, $\phi(\mu')$ is jointly concave in μ'. We next show that it is strictly concave. Note that since $\widehat{D}_t \ge 3(n+1)$, we can write

$$\mathbf{z}^\top \nabla^2_{\mu'} \phi(\mu') \mathbf{z} = -\sum_{i=0}^{3n+2} \frac{(\mathbf{z}^\top y^{i,t_{i+1}})^2}{({\mu'}^\top y^{i,t_{i+1}})^2} - \sum_{i=3(n+1)}^{\widehat{D}_t - 1} \frac{(\mathbf{z}^\top y^{i,t_{i+1}})^2}{({\mu'}^\top y^{i,t_{i+1}})^2}.$$

Since the columns of $\mathbf{Y}$ are linearly independent, then the first term in the right-hand side is strictly negative for any $\mathbf{z} \ne \mathbf{0}$. Therefore, $\mathbf{z}^\top \nabla^2_{\mu'} \phi(\mu') \mathbf{z} < 0$ for all $\mathbf{z} \ne \mathbf{0}$, hence $\phi(\mu')$ is strictly concave. □

A.2.18 Proof of Lemma II.3′

Proof. For simplicity of notation, we will use D_t instead of D_t^π to denote the cumulative adoptions at time t. Let $\mu = (\beta, \gamma)$ be the parameter vector where $\beta = (\beta_1, \beta_2, \beta_3)$ and $\gamma = (\gamma_j)_{j=0}^n$. From our discussion in Section II.5.4, note that the ML estimator $\hat{\mu}_t = (\hat{\beta}_t, \hat{\gamma}_t)$ is unique since $D_t \geq 3(n+1)$ and $\mathbf{Y}$ is full rank. Note from (II.4.2) that if either $\hat{\gamma}_j = +\infty$ or $\hat{\gamma}_j = -\infty$, then the likelihood function is 0 or negative. Then, we know there exist finite $\bar{\delta}_j, j = 0, 1, \cdots, n$ such that $\gamma_0 - \bar{\delta}_j \leq \hat{\gamma}_j \leq \gamma_0 + \bar{\delta}_j$.

If $\mu_0 = (\beta_0, \gamma_0)$ are the true parameters, note that $\hat{\mu}_t = (\hat{\beta}_t, \hat{\gamma}_t)$ can be written as:

$$\hat{\mu}_t = \arg \max_{\substack{\mu:\beta_1 \geq 0, \\ \beta_3 \leq 0}} \mathcal{L}_t(\widehat{\mathbf{U}}_t; \mu) = \mu_0 + \arg \min_{\substack{u:u_{b1} \geq -\beta_{01}, \\ u_{b3} \leq -\beta_{03}}} -\sum_{i=0}^{D_t} \ln \frac{f_i(\mu_0 + u)}{f_i(\mu_0)},$$

where $u = (u_{b1}, u_{b2}, u_{b3}, (u_{gj})_{j=0}^n)$. Let us denote by $\hat{u}$ the solution of the minimization problem on the right-hand side above. Hence, $\hat{\mu}_t = \mu_0 + \hat{u}$.

The result $\mathbb{E}_{\mu_0}[|\hat{\beta}_{tj} - \beta_{0j}|^2 \mid D_t = k] \leq \frac{\alpha_{\beta_j}}{k+1}$ for $j = 1, 2, 3$ and for some α_{β_j} that is independent of m_0 can be shown using a proof similar to that of Lemma II.3. Hence, to complete the proof of Lemma II.3′, we will need to establish that $\mathbb{E}_{\mu_0}[|\hat{\gamma}_{tj} - \gamma_{0j}|^2 \mid D_t = k] \leq \frac{\alpha_{\gamma_j}}{k+1}$ for all $j = 0, \ldots, n$ and for some α_{γ_j} independent of m_0.

We examine the estimation error $|\hat{\gamma}_{tj} - \gamma_{0j}|$ for some $j = 0, \ldots, n$. Let us denote e_j to be the $(n+4)$-dimensional binary vector, where the entry is equal to 1 only at the $(j+4)$-th index. Suppose that $|\hat{\gamma}_{tj} - \gamma_{0j}| > \delta$ for some $\bar{\delta}_j \geq \delta > 0$. This implies that $\hat{u}_{gj}$ lies outside the interval $[-\delta, \delta]$. Since the objective value is 0 when $u = 0$, and since the log-likelihood function is continuous and element-wise concave in γ_j, then either

$$-\sum_{i=0}^{D_t} \ln \frac{f_i(\mu_0 + \delta e_j)}{f_i(\mu_0)} \leq 0 \quad \text{or} \quad -\sum_{i=0}^{D_t} \ln \frac{f_i(\mu_0 - \delta e_j)}{f_i(\mu_0)} \leq 0.$$

Note that under the Markovian Bass model, the value $f_i(\mu)$ for any $\mu = (\beta, \gamma)$ is stochastic since its value depends on t_i and t_{i+1}, which are random adoption times. Here, t_i denotes the time of the i-th adoption, where $i = 0, \ldots, D_t$.

Let $\mathbb{P}_{\mu_0}(\cdot)$ denote the probability under a demand process that follows a Markovian Bass model with parameter vector μ_0. Therefore,

$$\begin{aligned} &\mathbb{P}_{\mu_0}\{|\hat{\gamma}_{tj} - \gamma_{0j}| > \delta\} \\ &\leq \mathbb{P}_{\mu_0}\left\{-\sum_{i=0}^{D_t} \ln \frac{f_i(\mu_0 + \delta e_j)}{f_i(\mu_0)} \leq 0\right\} + \mathbb{P}_{\mu_0}\left\{-\sum_{i=0}^{D_t} \ln \frac{f_i(\mu_0 - \delta e_j)}{f_i(\mu_0)} \leq 0\right\} \end{aligned}$$

$$= \mathbb{P}_{\mu_0}\left\{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}\geq 1\right\}+\mathbb{P}_{\mu_0}\left\{\prod_{i=0}^{D_t}\frac{f_i(\mu_0-\delta e_j)}{f_i(\mu_0)}\geq 1\right\}$$

$$\leq \mathbb{P}_{\mu_0}\left\{\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\geq 1\right\}+\mathbb{P}_{\mu_0}\left\{\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0-\delta e_j)}{f_i(\mu_0)}}\geq 1\right\}$$

$$\leq \mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\right)+\mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0-\delta e_j)}{f_i(\mu_0)}}\right). \tag{A.92}$$

Hence, we need to bound the two terms in (A.92). We demonstrate how we can bound the first term, since the second term can be bounded following similar arguments. By the law of iterated expectations we know that the first term in (A.92) can be written as

$$\mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\right)=\mathbb{E}_{\mu_0}\left(\cdots\mathbb{E}_{\mu_0}\left(\mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)\mid\mathcal{F}_{t_{D_t-2}}\right)\cdots\mid\mathcal{F}_0\right). \tag{A.93}$$

We will analyze this expression, starting from the innermost conditional expectation.

Note that

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\mathbb{E}_{\mu_0}\left(\sqrt{\frac{f_{D_t}(\mu_0+\delta e_j)}{f_{D_t}(\mu_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)$$

$$=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\left(\int_{t_{D_t-1}}^{\infty}\sqrt{\frac{f_{D_t}(\mu_0+\delta e_j)}{f_{D_t}(\mu_0)}}f_{D_t}(\mu_0)\mathrm{d}t_{D_t}\right)$$

$$=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\left(\int_{t_{D_t-1}}^{\infty}\sqrt{f_{D_t}(\mu_0+\delta e_j)}\sqrt{f_{D_t}(\mu_0)}\mathrm{d}t_{D_t}\right). \tag{A.94}$$

Here, the first equality is because $\{f_i(\mu), i=0,\ldots,D_t-1\}$ are all $\mathcal{F}_{t_{D_t-1}}$-measurable. The second equality is because, given the information set $\mathcal{F}_{t_{D_t-1}}$, $f_{D_t}(\mu_0)$ is the conditional probability density function of the adoption time t_{D_t} under a Markovian Bass model with parameter vector μ_0. Hence, we will next derive a bound on

$$\int_{t_{D_t-1}}^{\infty}\sqrt{f_{D_t}(\mu_0+\delta e_j)}\sqrt{f_{D_t}(\mu_0)}\mathrm{d}t_{D_t}.$$

Note that

$$
\begin{aligned}
&\frac{1}{2}\int_{t_{D_{t-1}}}^{\infty}\left(\sqrt{f_{D_t}(\mu_0+\delta e_j)}-\sqrt{f_{D_t}(\mu_0)}\right)^2 \mathrm{d}t_{D_t} \\
&=\frac{1}{2}\int_{t_{D_{t-1}}}^{\infty}\left(f_{D_t}(\mu_0+\delta e_j)+f_{D_t}(\mu_0)-2\sqrt{f_{D_t}(\mu_0+\delta e_j)f_{D_t}(\mu_0)}\right)\mathrm{d}t_{D_t} \\
&=1-\int_{t_{D_{t-1}}}^{\infty}\sqrt{f_{D_t}(\mu_0+\delta e_j)f_{D_t}(\mu_0)}\mathrm{d}t_{D_t},
\end{aligned}
$$

where the last equality is because the integral of the probability density function

$$
\int_{t_{D_{t-1}}}^{\infty} f_{D_t}(\mu)\mathrm{d}t_{D_t}
$$

is equal to 1 for any μ. Therefore,

$$
\int_{t_{D_{t-1}}}^{\infty}\sqrt{f_{D_t}(\mu_0+\delta e_j)f_{D_t}(\mu_0)}\mathrm{d}t_{D_t}=1-\frac{1}{2}\int_{t_{D_{t-1}}}^{\infty}\left(\sqrt{f_{D_t}(\mu_0+\delta e_j)}-\sqrt{f_{D_t}(\mu_0)}\right)^2\mathrm{d}t_{D_t}. \tag{A.95}
$$

The integral on the right-hand side is the Hellinger distance between $f_{D_t}(\mu_0+\delta e_j)$ and $f_{D_t}(\mu_0)$, which are probability densities of the adoption time t_{D_t}.

Note that the Hellinger distance can be lower bounded by the K-L divergence (see corollary 4.9 in Taneja and Kumar 2004). Specifically,

$$
\frac{1}{2}\int_{t_{D_{t-1}}}^{\infty}\left(\sqrt{f_{D_t}(\mu_0+\delta e_j)}-\sqrt{f_{D_t}(\mu_0)}\right)^2 \geq \frac{1}{4\sqrt{R}}\mathbb{E}_{\mu_0}\left(\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\mu_0+\delta e_j)}\mid \mathcal{F}_{t_{D_{t-1}}}\right), \tag{A.96}
$$

where R is a constant such that $R\geq \min_\delta \frac{1}{f_{D_t}(\mu_0+\delta e_j)}\geq\frac{1}{m_0p_0\bar{x}^u}$, where $\bar{x}^u$ is defined in Assumption II.1. We will next derive a bound on the right-hand side.

Note that if we define $C_I := (x(r;\gamma_0)+\bar{\delta}_j b_{j,n}(r))^2/b_{j,n}(r)^2$ for some $r\in(0,1)$, we have

$$
\frac{\partial}{\partial\delta}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\mu_0+\delta e_j)}=-\frac{b_{j,n}(r_{t_{D_t}})}{\sum_{i\neq j}\gamma_{0i}b_{i,n}(r_{t_{D_t}})+(\gamma_{0j}+\delta)b_{j,n}(r_{t_{D_t}})},
$$

and

$$\frac{\partial^2}{\partial\delta^2}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\mu_0+\delta e_j)} = \frac{b_{j,n}(r_{t_{D_t}})^2}{\left(\sum_{i\neq j}\gamma_{0i}b_{i,n}(r_{t_{D_t}})+(\gamma_{0j}+\delta)b_{j,n}(r_{t_{D_t}})\right)^2} \geq \frac{1}{C_I}.$$

Note that C_I is independent of m_0.

Furthermore, since the expectation of the Fisher score under the true parameter is zero, we have

$$\mathbb{E}_{\mu_0}\left(\frac{\partial}{\partial\delta}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\mu_0+\delta e_j)}\Big|_{\delta=0} \mid \mathcal{F}_{t_{D_t-1}}\right)=0.$$

Hence, a simple calculation yields

$$\begin{aligned}
\mathbb{E}_{\mu_0}\left(\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\mu_0+\delta e_j)} \mid \mathcal{F}_{t_{D_t-1}}\right) &= \mathbb{E}_{\mu_0}\left(\int_0^\delta \frac{\partial}{\partial z}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\mu_0+z e_j)}\mathrm{d}z \mid \mathcal{F}_{t_{D_t-1}}\right)\\
&= \mathbb{E}_{\mu_0}\left(\int_0^\delta\left(\frac{\partial}{\partial z}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\theta_0+z e_j)}-\frac{\partial}{\partial z}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\theta_0+z e_j)}\Big|_{z=0}\right)\mathrm{d}z \mid \mathcal{F}_{t_{D_t-1}}\right)\\
&= \mathbb{E}_{\mu_0}\left(\int_0^\delta\int_0^z \frac{\partial^2}{\partial z'^2}\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\mu_0+z' e_j)}\mathrm{d}z' \mid \mathcal{F}_{t_{D_t-1}}\right) \geq \frac{1}{2C_I}\delta^2.
\end{aligned}$$

Then, (A.96) reduces to

$$\frac{1}{4\sqrt{R}C_I}\delta^2 \leq \int_{t_{D_t-1}}^\infty\left(\sqrt{f_{D_t}(\mu_0+\delta e_j)}-\sqrt{f_{D_t}(\mu_0)}\right)^2\mathrm{d}t_{D_t-1}.$$

Hence, from (A.95), we have

$$\begin{aligned}
\int_{t_{D_t-1}}^\infty\sqrt{f_{D_t}(\mu_0+\delta e_j)f_{D_t}(\mu_0)}\mathrm{d}t_{D_t} &= 1-\frac{1}{2}\int_{t_{D_t-1}}^\infty\left(\sqrt{f_{D_t}(\mu_0+\delta e_j)}-\sqrt{f_{D_t}(\mu_0)}\right)^2\mathrm{d}t_{D_t}\\
&\leq \exp\left(-\frac{1}{2}\int_{t_{D_t-1}}^\infty\left(\sqrt{f_{D_t}(\mu_0+\delta e_j)}-\sqrt{f_{D_t}(\mu_0)}\right)^2\mathrm{d}t_{D_t}\right) \leq \exp\left(-\frac{1}{8\sqrt{R}C_I}\delta^2\right),
\end{aligned}$$

where the first inequality is because $e^{-x}\geq 1-x$ for all x.

Hence, from (A.94), we have

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}} \mid \mathcal{F}_{t_{D_t-1}}\right) \leq \sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\exp\left(-\frac{1}{8\sqrt{R}C_I}\delta^2\right). \quad \text{(A.97)}$$

This provides a bound for the innermost conditional expectation in (A.93). Observe that all of the terms in the right-hand side of (A.97) are $\mathcal{F}_{t_{D_t-2}}$-measurable, except for the term $\sqrt{f_{D_t-1}(\mu_0+\delta e_j)/f_{D_t-1}(\mu_0)}$. Hence, if we take the conditional expectation on both sides of (A.97) given $\mathcal{F}_{t_{D_t-2}}$, and use the same logic as our arguments above, we get

$$\mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}} \mid \mathcal{F}_{t_{D_t-2}}\right) \leq \sqrt{\prod_{i=0}^{D_t-2}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\cdot\exp\left(-\frac{2}{8\sqrt{R}C_I}\delta^2\right)$$

We can proceed iteratively to evaluate (A.94) as we take conditional expectations given $\mathcal{F}_{t_{D_t}-3}$, $\mathcal{F}_{t_{D_t}-4}$, $\mathcal{F}_0$, resulting in

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\right) \leq \mathbb{E}_{\mu_0}\left(\exp\left(-\frac{D_t+1}{8\sqrt{R}C_I}\delta^2\right)\right)$$

Using similar arguments, we can get the same bound for the second term in (A.92). Therefore, we have

$$\mathbb{P}_{\mu_0}\{|\hat{\gamma}_{tj}-\gamma_{0j}|>\delta\} \leq 2\mathbb{E}_{\mu_0}\left(\exp\left(-\frac{D_t+1}{8\sqrt{R}C_I}\delta^2\right)\right).$$

Hence,

$$\begin{aligned}
\mathbb{E}_{\mu_0}\left[(\hat{\gamma}_{tj}-\gamma_{0j})^2 \mid D_t=k\right] &= \int_0^\infty \mathbb{P}_{\mu_0}\left\{(\hat{\gamma}_{tj}-\gamma_{0j})^2>\delta \mid D_t=k\right\}\mathrm{d}\delta \\
&= \int_0^\infty \mathbb{P}_{\mu_0}\left\{|\hat{\gamma}_{tj}-\gamma_{0j}|^2>\sqrt{\delta} \mid D_t=k\right\}\mathrm{d}\delta \\
&\leq \int_0^\infty \exp\left(-\frac{k+1}{8\sqrt{R}C_I}\delta\right)\mathrm{d}\delta = \frac{8\sqrt{R}C_I}{k+1}.
\end{aligned}$$

□

A.2.19 Proof of Lemma II.4′

Proof. The proof of Lemma II.4′ follows exactly the same steps as the proof of Lemma II.4. The only difference is to show $\left|\frac{\partial r_t^*}{\partial \gamma_i}\right|$ is bounded. Because $x(r)$ is linear in every γ_i, we have γ_i to be in the similar positions to p or q. Therefore, following the steps to bound $\left|\frac{\partial r_t^*}{\partial p}\right|$ or $\left|\frac{\partial r_t^*}{\partial q}\right|$ gives us the desired result. We next discuss how to bound $\left|\frac{\partial r_t^*}{\partial \gamma_i}\right|$.

Using the equation (II.3.2) satisfied by $r_t^*(\theta, d)$, we differentiate r_t^* with respect to γ_i and rearranging terms, we get that for any d,

$$\left|\frac{\partial r_t^*(\mu, d)}{\partial \gamma_i}\right| = \left|\frac{\frac{\partial}{\partial \gamma_i}[V(d, T-t) - V(d+1, T-t)] - \frac{b_{i,n}(r_t^*)x'(r_t^*) - x(r_t^*)\binom{n}{i} r_t^{*\,i-1}(1-r_t^*)^{n-i-1}(i - nr_t^*)}{x'(r_t^*)^2}}{\frac{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}{x'(r_t^*)^2}}\right|. \tag{A.98}$$

Similar to (A.67), we have

$$\underbrace{\frac{\partial}{\partial \gamma_i}[V(d, T-t) - V(d+1, T-t)]}_{\text{(A1)}} = \underbrace{\frac{\partial^2 V(d, T-t)}{\partial \gamma_i \partial t} \cdot \frac{1}{(m-d)\left(p + \frac{d}{m}q\right)x(r_t^*)}}_{\text{(A2)}} - \frac{2x(r_t^*)x'(r_t^*)b_{i,n}(r_t^*) - x(r_t^*)^2\binom{n}{i} r_t^{*\,i-1}(1-r_t^*)^{n-i-1}(i - nr_t^*)}{x'(r_t^*)^2}. \tag{A.99}$$

We now examine the absolute value of (A2) on the right-hand side of (A.99). Because the partial effect of γ_i on the expected revenue rate cannot exceed the rate when all the remaining population $(m-d)$ are directly affected by γ_i without being affected by the current price r_t^*, we know

$$\left|\frac{\partial^2 V(d, T-t)}{\partial \gamma_i \partial t}\right| \le (m-d)(p + qd/m)r_t^* b_{i,n}(r_t^*).$$

Hence, we can bound the absolute value of (A1) as follows:

$$|\text{(A1)}| \le \frac{r_t^* b_{i,n}(r_t^*)}{x(r_t^*)} + \left|\frac{2x(r_t^*)x'(r_t^*)b_{i,n}(r_t^*) - x(r_t^*)^2\binom{n}{i} r_t^{*\,i-1}(1-r_t^*)^{n-i-1}(i - nr_t^*)}{x'(r_t^*)^2}\right|. \tag{A.100}$$

Therefore, we substitute (A.100) into (A.98) to get

$$\begin{aligned}\left|\frac{\partial r_t^*(\mu,d)}{\partial \gamma_i}\right| &\leq \left|\frac{r_t^* b_{i,n}(r_t^*) x'(r_t^*)^2}{x(r_t^*)\left(2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)\right)}\right| \\ &+ \left|\frac{2x(r_t^*)x'(r_t^*)b_{i,n}(r_t^*) - x(r_t^*)^2\binom{n}{i}r_t^{*i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right| \\ &+ \left|\frac{b_{i,n}(r_t^*)x'(r_t^*) - x(r_t^*)\binom{n}{i}r_t^{*i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right| \\ &\leq \frac{M^2}{\bar{x}^l C_d} b_{i,n}(r_t^*) + \frac{2M\bar{x}^u}{C_d} b_{i,n}(r_t^*) + \left|\frac{\bar{x}^{u2}\binom{n}{i}r_t^{*i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{C_d}\right| + \frac{M}{C_d} b_{i,n}(r_t^*) \\ &+ \left|\frac{\bar{x}^u\binom{n}{i}r_t^{*i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{C_d}\right|\end{aligned}$$

where the inequality follows from Assumption II.1(i),(ii), (iv). Note that all terms on the RHS of the inequality does not scale up with the market size m_0 and is finite since r_t^* does not scale up with the market size m_0 all γ_i are finite.

□

A.2.20 Proof of Proposition A.1

Proof. Suppose the value of m_0 is known. Given the data $\widehat{\mathbf{U}}_t = \{(r_s^\pi, D_s^\pi), s \leq t\}$ at time $t \geq t_2$, let $\hat{\theta}_t = (\hat{p}_t, \hat{q}_t)$ denote the ML estimator of (p_0, q_0). Note that since $D_t^\pi \geq 2$, we can verify that $\mathcal{L}_t$ in (II.4.3) is strongly concave. Hence, $(\hat{p}_t, \hat{q}_t)$ is the unique maximizer of $\mathcal{L}_t$. Since the log-likelihood function is jointly concave in β_t (Proposition II.2), and since β_t is an affine function of $\hat{\theta}_t$, then $\hat{\theta}_t$ satisfy first-order conditions. Specifically, $(\hat{p}_t, \hat{q}_t)$ is the solution to the following system of equations in variables (p, q):

$$0 = \frac{\partial \mathcal{L}_t}{\partial p} = \sum_{j=0}^{D_t^\pi - 1} \frac{1}{p + \frac{j}{m_0} q} - \int_0^t (m_0 - D_s^\pi) x(r_s^\pi) \mathrm{d}s, \tag{A.101}$$

$$0 = \frac{\partial \mathcal{L}_t}{\partial q} = \sum_{j=0}^{D_t^\pi - 1} \frac{\frac{j}{m_0}}{p + \frac{j}{m_0} q} - \int_0^t (m_0 - D_s^\pi) \frac{D_s^\pi}{m_0} x(r_s^\pi) \mathrm{d}s. \tag{A.102}$$

Note that the value of the ML estimator $(\hat{p}_t, \hat{q}_t)$ is changing continuously in time. This is because the first-order conditions (A.101),(A.102) are changing over time. The temporal change of the first-order conditions is impacted by a combination of a continuous phenomenon (time t) and a discrete phenomenon (the pure jump process D_t^π). Hence, the differential $(\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t)$ is also impacted by a continuous phenomenon and a discrete phenomenon. However, an explicit form of the differential equation is not immediately obvious from (A.101),(A.102).

We will apply Itô's lemma to derive the stochastic differential equation of $(\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t)$. Intuitively speaking, Itô's lemma is a chain rule defined on stochastic processes. Many readers may be more familiar with Itô's lemma as applied to functions of Brownian motion. Here, our call is to a modified Itô's lemma that applies to jump processes (see Section 11.5.1 in *Shreve* 2004). Specifically, suppose that $X_t = X_t^c + D_t$ is a stochastic process, where D_t is a pure jump process and X_t^c is a continuous-path process with differential form:

$$X_t^c = X_0^c + \int_0^t \Gamma_s \mathrm{d}W_s + \int_0^t \Theta_s \mathrm{d}s,$$

where W_t is a standard Wiener process (Brownian motion) and Γ_s and Θ_s are adapted processes. In differential notation, we write

$$\mathrm{d}X_s^c = \Gamma_s \mathrm{d}W_s + \Theta_s \mathrm{d}s, \qquad \mathrm{d}X_s^c \mathrm{d}X_s^c = \Gamma_s^2 \mathrm{d}s.$$

Then the following theorem (adapted from Theorem 11.5.4 in *Shreve* 2004) provides the expression for the dynamics of a function of a jump process X_t.

Theorem A.1 (Itô-Doeblin formula for a jump process)**.** Let X_t be a jump process and let $f(t, x)$ a function whose first and second partial derivatives appearing in the following formula are defined and are continuous. Then,

$$f(t, X_t) = f(0, X_0) + \int_0^t f_t(s, X_s)\mathrm{d}s + \int_0^t f_x(s, X_s)\mathrm{d}X_s^c + \frac{1}{2}\int_0^t f_{xx}(s, X_s)\mathrm{d}X_s^c \mathrm{d}X_s^c$$
$$+ \sum_{0<s\leq t} \left[f(s, X_s) - f(s, X_{s-}) \right].$$

Let $f^p(t, D_t^\pi)$, $f^q(t, D_t^\pi)$ denote the functions that map (t, D_t^π) to the ML estimators $\hat{p}_t$, $\hat{q}_t$. Note that these functions are defined from the implicit equations (A.101),(A.102). Hence, since $(\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t) = (\mathrm{d}f^p(t, D_t^\pi), \mathrm{d}f^q(t, D_t^\pi))$, we can get the differential equation for the ML estimators by applying Itô's lemma on the functions f^p, f^q of the pure jump process D_t^π.

Using the notation from the Itô-Doeblin formula, $X_t = X_t^c + D_t^\pi$ in our case is a jump process where $X_t^c = 0$ for all t. Hence, from Theorem A.1, any function $f(t, D_t^\pi)$ satisfies

$$f(t, D_t^\pi) = f(0, D_0^\pi) + \int_0^t f_t(s, D_s^\pi)\mathrm{d}s + \sum_{0<s\leq t} \left[f(s, D_s^\pi) - f(s, D_{s-}^\pi)\right].$$

We can write this in differential form as:

$$\mathrm{d}f(t, D_t^\pi) = f_t(t, D_t^\pi)\mathrm{d}t + \left[f(t, D_{t-}^\pi + 1) - f(t, D_{t-}^\pi)\right]\mathrm{d}D_t^\pi.$$

We can apply this result to get the differential equations of the ML estimators:

$$\mathrm{d}\hat{p}_t = \mathrm{d}f^p(t, D_t^\pi) = \underbrace{f_t^p(t, D_t^\pi)}_{=\mu_p}\mathrm{d}t + \underbrace{\left[f^p(t, D_{t-}^\pi + 1) - f^p(t, D_{t-}^\pi)\right]}_{=\eta_p}\mathrm{d}D_t^\pi \tag{A.103}$$

$$\mathrm{d}\hat{q}_t = \mathrm{d}f^q(t, D_t^\pi) = \underbrace{f_t^q(t, D_t^\pi)}_{=\mu_q}\mathrm{d}t + \underbrace{\left[f^q(t, D_{t-}^\pi + 1) - f^q(t, D_{t-}^\pi)\right]}_{=\eta_q}\mathrm{d}D_t^\pi. \tag{A.104}$$

Therefore, in order to determine the explicit form of the differential equations $(\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t)$, we need to determine the continuous change (μ_p and μ_q), as well as the discrete change (η_p and η_q).

As can be seen in (A.103),(A.104), the continuous change μ_p,μ_q in the ML estimators can be found by deriving the expressions for $f_t^p = \frac{\partial \hat{p}_t}{\partial t}$ and $f_t^q = \frac{\partial \hat{q}_t}{\partial t}$. Recall that $(\hat{p}_t, \hat{q}_t)$ are the solutions to the system of equations (A.101),(A.102). Therefore, we can substitute $(\hat{p}_t, \hat{q}_t)$ into these equations, then take their partial derivative with respect to t. This gives us:

$$0 = -\left(\sum_{j=0}^{D_t^\pi - 1} \frac{1}{\left(\hat{p}_t + \frac{j}{m_0}\hat{q}_t\right)^2}\right)\frac{\partial \hat{p}_t}{\partial t} - \left(\sum_{j=0}^{D_t^\pi - 1} \frac{\frac{j}{m_0}}{\left(\hat{p}_t + \frac{j}{m_0}\hat{q}_t\right)^2}\right)\frac{\partial \hat{q}_t}{\partial t} - (m_0 - D_t^\pi)x(r_t^\pi),$$

$$0 = -\left(\sum_{j=0}^{D_t^\pi - 1} \frac{\frac{j}{m_0}}{\left(\hat{p}_t + \frac{j}{m_0}\hat{q}_t\right)^2}\right)\frac{\partial \hat{p}_t}{\partial t} - \left(\sum_{j=0}^{D_t^\pi - 1} \frac{\left(\frac{j}{m_0}\right)^2}{\left(\hat{p}_t + \frac{j}{m_0}\hat{q}_t\right)^2}\right)\frac{\partial \hat{q}_t}{\partial t} - (m_0 - D_t^\pi)\frac{D_t^\pi}{m_0}x(r_t^\pi).$$

Solving the system of equations for $\frac{\partial \hat{p}_t}{\partial t}$ and $\frac{\partial \hat{q}_t}{\partial t}$, we get

$$\mu_p = f_t^p(t, D_t^\pi) = \frac{\partial \hat{p}_t}{\partial t} = -\left(\sigma_p^2 + \sigma_{pq}^2\frac{D_t^\pi}{m_0}\right)(m_0 - D_t^\pi)x(r_t^\pi), \tag{A.105}$$

$$\mu_q = f_t^q(t, D_t^\pi) = \frac{\partial \hat{q}_t}{\partial t} = -\left(\sigma_{pq}^2 + \sigma_q^2\frac{D_t^\pi}{m_0}\right)(m_0 - D_t^\pi)x(r_t^\pi) \tag{A.106}$$

where we define

$$\begin{pmatrix} \sigma_p^2 & \sigma_{pq}^2 \\ \sigma_{pq}^2 & \sigma_q^2 \end{pmatrix} := \begin{pmatrix} \frac{1}{\Delta_2}\sum_{j=0}^{D_t^\pi-1}\frac{\left(\frac{j}{m_0}\right)^2}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2} & -\frac{1}{\Delta_2}\sum_{j=0}^{D_t^\pi-1}\frac{\frac{j}{m_0}}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2} \\ -\frac{1}{\Delta_2}\sum_{j=0}^{D_t^\pi-1}\frac{\frac{j}{m_0}}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2} & \frac{1}{\Delta_2}\sum_{j=0}^{D_t^\pi-1}\frac{1}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2} \end{pmatrix} \tag{A.107}$$

and

$$\Delta_2 := \left(\sum_{j=0}^{D_t^\pi-1}\frac{1}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2}\right)\left(\sum_{j=0}^{D_t^\pi-1}\frac{\left(\frac{j}{m_0}\right)^2}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2}\right) - \left(\sum_{j=0}^{D_t^\pi-1}\frac{\frac{j}{m_0}}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2}\right)^2.$$

Note that $\Delta_2 > 0$ by Cauchy-Schwarz inequality, so all elements in the matrix (A.107) are finite. Hence, μ_p and μ_q defined in (A.105),(A.106) are the terms multiplied to $\mathrm{d}t$ in the stochastic differential equations of $\mathrm{d}\hat{p}_t$, $\mathrm{d}\hat{q}_t$, respectively, defined in (A.103),(A.104).

We can further simplify the expressions for $\sigma_p^2, \sigma_{pq}^2, \sigma_q^2$. We do this by first defining the vectors $\mathbf{v}_p = (v_{p,j}), \mathbf{v}_q = (v_{q,j})$ to be D_t^π-dimensional vectors with entries

$$v_{p,j} = \frac{1}{\hat{p}_t + \frac{j-1}{m_0}\hat{q}_t}, \quad v_{q,j} = \frac{\frac{j-1}{m_0}}{\hat{p}_t + \frac{j-1}{m_0}\hat{q}_t}, \qquad \text{for } j = 1, \ldots, D_t^\pi$$

If α is the angle between vectors $\mathbf{a}$ and $\mathbf{b}$, then due to the identity $\cos\alpha = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|}$, we have that

$$1 - \cos^2\alpha = \frac{\Delta_2}{\left(\sum_{j=0}^{D_t^\pi-1}\frac{1}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2}\right)\left(\sum_{j=0}^{D_t^\pi-1}\frac{\left(\frac{j}{m_0}\right)^2}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2}\right)}.$$

Therefore, we can rewrite $\sigma_p^2, \sigma_{pq}^2, \sigma_q^2$ as

$$\sigma_p^2 = \frac{1}{1-\cos^2\alpha}\left(\sum_{j=0}^{D_t^\pi-1}\frac{1}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2}\right)^{-1} = \frac{1}{1-\cos^2\alpha}\cdot\frac{1}{\|\mathbf{v}_p\|^2}$$

$$\sigma_q^2 = \frac{1}{1-\cos^2\alpha}\left(\sum_{j=0}^{D_t^\pi-1}\frac{\left(\frac{j}{m_0}\right)^2}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2}\right)^{-1} = \frac{1}{1-\cos^2\alpha}\cdot\frac{1}{\|\mathbf{v}_q\|^2}$$

$$\sigma_{pq}^2 = -\frac{\cos^2\alpha}{1-\cos^2\alpha}\left(\sum_{j=0}^{D_t^\pi-1}\frac{\frac{j}{m_0}}{\left(\hat{p}_t+\frac{j}{m_0}\hat{q}_t\right)^2}\right)^{-1} = -\frac{\cos^2\alpha}{1-\cos^2\alpha}\cdot\frac{1}{\mathbf{v}_p^\top\mathbf{v}_q}$$

We next derive the expression for the jump size η_p, η_q in (A.103),(A.104). We will use the property that $\mathbb{E}_{\theta_0}(\hat{p}_t \mid \mathcal{F}_t) = \hat{p}_t$ and $\mathbb{E}_{\theta_0}(\hat{q}_t \mid \mathcal{F}_t) = \hat{q}_t$. Therefore, $\hat{p}_t$ and $\hat{q}_t$ are $\mathcal{F}_t$-martingales, by definition, with $\mathbb{E}_{\theta_0}(\mathrm{d}\hat{p}_t \mid \mathcal{F}_t) = \mathbb{E}_{\theta_0}(\mathrm{d}\hat{q}_t \mid \mathcal{F}_t) = 0$. Therefore, taking the conditional expectation of (A.103), (A.104), and using the property that $\mathbb{E}_{\theta_0}(\mathrm{d}D_t^\pi \mid \mathcal{F}_t) = (m_0 - D_t^\pi)\left(\hat{p}_t + \frac{D_t^\pi}{m_0}\hat{q}_t\right)x(r_t^\pi)\mathrm{d}t$, we have

$$\begin{aligned}
0 = \mathbb{E}_{\theta_0}(\mathrm{d}\hat{p}_t \mid \mathcal{F}_t) &= \left[\mu_p + \eta_p(m_0 - D_t^\pi)\left(\hat{p}_t + \frac{D_t^\pi}{m_0}\hat{q}_t\right)x(r_t^\pi)\right]\mathrm{d}t \\
0 = \mathbb{E}_{\theta_0}(\mathrm{d}\hat{q}_t \mid \mathcal{F}_t) &= \left[\mu_q + \eta_q(m_0 - D_t^\pi)\left(\hat{p}_t + \frac{D_t^\pi}{m_0}\hat{q}_t\right)x(r_t^\pi)\right]\mathrm{d}t.
\end{aligned}$$

Using the expression of μ_p, μ_q from (A.105),(A.106), we can solve for η_p, η_q:

$$\eta_p = \frac{-\mu_p}{(m_0 - D_t^\pi)\left(\hat{p}_t + \frac{D_t^\pi}{m_0}\hat{q}_t\right)x(r_t^\pi)} = \frac{\sigma_p^2 + \sigma_{pq}^2\frac{D_t^\pi}{m_0}}{\hat{p}_t + \frac{D_t^\pi}{m_0}\hat{q}_t}, \tag{A.108}$$

$$\eta_q = \frac{-\mu_q}{(m_0 - D_t^\pi)\left(\hat{p}_t + \frac{D_t^\pi}{m_0}\hat{q}_t\right)x(r_t^\pi)} = \frac{\sigma_{pq}^2 + \sigma_q^2\frac{D_t^\pi}{m_0}}{\hat{p}_t + \frac{D_t^\pi}{m_0}\hat{q}_t}. \tag{A.109}$$

Substituting the expressions (A.105),(A.106) for μ_p, μ_q and expressions (A.108),(A.109) for η_p, η_q into (A.103), (A.104) results in the differential equations (A.4) and (A.5). This concludes the proof of the proposition. □

A.2.21 Proof of Theorem A.2

Proof. The dynamic programming formulation is

$$\begin{aligned}
&V^{\mathsf{MLE}}(d,t,p,q) \\
&= \max_r \Big\{ r\lambda(d,r)\delta t + V^{\mathsf{MLE}}(d+1, t-\delta t, p+\eta_p+\mu_p(r)\delta t, q+\eta_q+\mu_q(r)\delta t)\lambda(d,r)\delta t \\
&+ V^{\mathsf{MLE}}(d, t-\delta t, p+\mu_p(r)\delta t, q+\mu_q(r)\delta t)\left[1-\lambda(d,r)\delta t\right] + o(\delta t)\Big\}.
\end{aligned}$$

Subtracting both sides by $V^{\mathsf{MLE}}(d, t-\delta t)$ and then dividing by δt results in

$$
\begin{aligned}
&\frac{1}{\delta t}\left[V^{\mathsf{MLE}}(d,t,p,q) - V^{\mathsf{MLE}}(d,t-\delta t,p,q)\right] \\
&= \max_r \Bigg\{ r\lambda(d,r) + \left[V^{\mathsf{MLE}}(d+1,t-\delta t,p+\eta_p,q+\eta_q) - V^{\mathsf{MLE}}(d,t-\delta t,p+\mu_p(r)\delta t,q+\mu_q(r)\delta t)\right]\cdot\lambda(d,r) \\
&\qquad + \frac{V^{\mathsf{MLE}}(d,t-\delta t,p+\mu_p(r)\delta t,q+\mu_q(r)\delta t) - V^{\mathsf{MLE}}(d,t-\delta t,p,q+\mu_q(r)\delta t)}{\mu_p(r)\delta t}\cdot\mu_p(r) \\
&\qquad + \frac{V^{\mathsf{MLE}}(d,t-\delta t,p,q+\mu_q(r)\delta t) - V^{\mathsf{MLE}}(d,t-\delta t,p,q)}{\mu_q(r)\delta t}\cdot\mu_q(r) + \frac{o(\delta t)}{\delta t}\Bigg\},
\end{aligned}
$$

where $\mu_p(r), \mu_q(r), \eta_p, \eta_q$ are defined in (A.8),(A.9). Substituting the expressions for $\mu_p(r)$,$\mu_q(r)$, then taking the limit on both sides as δt approaches 0, and using (A.4),(A.5), we get

$$
\begin{aligned}
\frac{\partial V^{\mathsf{MLE}}}{\partial t} = \max_r \quad &\Bigg\{ r\lambda(d,r) + \left[V^{\mathsf{MLE}}(d+1,t,p+\eta_p,q+\eta_q) - V^{\mathsf{MLE}}(d,t,p,q)\right]\lambda(d,r) \\
&- \frac{\partial V^{\mathsf{MLE}}}{\partial p}\left(\sigma_p^2 + \sigma_{pq}^2\frac{d}{m_0}\right)(m_0-d)x(r) - \frac{\partial V^{\mathsf{MLE}}}{\partial q}\left(\sigma_{pq}^2 + \sigma_q^2\frac{d}{m_0}\right)(m_0-d)x(r)\Bigg\}.
\end{aligned}
$$

We denote the objective function of RHS by $J^{\mathsf{MLE}}(d,t,p,q,r)$. Note that

$$
\begin{aligned}
\frac{\partial J^{\mathsf{MLE}}}{\partial r} &= \frac{\lambda(d,r)}{x(r)}\left(rx'(r) + x(r)\right) - \left[\frac{\partial V^{\mathsf{MLE}}}{\partial p}\left(\sigma_p^2 + \sigma_{pq}^2\frac{d}{m_0}\right) + \frac{\partial V^{\mathsf{MLE}}}{\partial q}\left(\sigma_{pq}^2 + \sigma_q^2\frac{d}{m_0}\right)\right](m_0-d)x'(r) \\
&\quad + \lambda(d,r)\frac{x'(r)}{x(r)}\left[V(d+1,t,p+\eta_p,q+\eta_q) - V(d,t,p,q)\right]
\end{aligned}
$$

We can verify that the first order sufficient condition $\frac{\partial J^{\mathsf{MLE}}}{\partial r} = 0$ is satisfied by $r^*(d,t,p,q)$ in (A.10). Plugging back r^* to the dynamic program, we get (A.11). □

A.3 Pricing and learning in a Markovian Bass model

Suppose that the seller knows the demand follows a Markovian Bass model, however she does not know the true parameter vector θ_0. Instead, at time t she infers the unknown parameters from the information set $\mathcal{F}_t$. Specifically, she infers the parameters from the price and sales data $\widehat{\mathbf{U}}_t$ at time t using a statistical inference method.

Recall the seller's pricing-and-learning problem:

$$
\sup_{\pi\in\Pi} \mathbb{E}_{D|\mathcal{F}_0}\left[\int_0^T r_t^\pi \, \mathrm{d}D_t^\pi \;\Big|\; \mathcal{F}_0\right] \;=\; \sup_{\pi\in\Pi} \mathbb{E}_{D|\mathcal{F}_0}\left[\int_0^T \mathbb{E}_{D|\mathcal{F}_t}\left[r_t^\pi \, \mathrm{d}D_t^\pi \mid \mathcal{F}_t\right] \;\Big|\; \mathcal{F}_0\right]. \tag{A.1}
$$

Here, her objective is to choose a $\mathcal{F}_t$-adapted pricing policy that maximizes her total expected revenue. Since she does not know θ_0, the conditional expectation $\mathbb{E}_{D|\mathcal{F}_t}$ in (A.1) is with respect to the inferred demand distribution from the price and sales data at time t.

In general, there are two schools of statistical inference: Bayesian and frequentist. A Bayesian framework represents uncertainty about the possible values of the demand parameters through a probability distribution. The evidence (data) collected over time is used to update the confidence about the inference, represented by a posterior probability distribution over the possible parameter values. If the seller uses Bayesian inference to learn the demand parameters, let $V^{\mathsf{B}}(t, \widehat{\mathbf{U}}_t)$ denote her inferred optimal expected revenue-to-go at time t given data $\widehat{\mathbf{U}}_t = ((r_u, D_u), u \leq t)$. We can write V^{B} as:

$$
\begin{aligned}
V^{\mathsf{B}}(t, \widehat{\mathbf{U}}_t) &= \max_{r_t} \left\{ \sum_{\theta \in \Theta} \mathbb{P}(\theta \mid \widehat{\mathbf{U}}_t) \cdot \mathbb{E}_{\theta} \left[r_t \cdot (D_{t+h} - D_t) + V^{\mathsf{B}} \left(t+h, \widehat{\mathbf{U}}_t \cup (r_t, D_{t+h}) \right) \;\middle|\; D_t \right] \right\} \qquad \text{(A.2)} \\
&= \max_{r_t} \left\{ \sum_{\theta \in \Theta} \mathbb{P}(\theta \mid \widehat{\mathbf{U}}_t) \cdot \Big\{ \lambda(D_t, r_t; \theta) h \cdot \left[r_t + V^{\mathsf{B}} \left(t+h, \widehat{\mathbf{U}}_t \cup (r_t, D_t + 1) \right) \right] \right. \\
&\qquad \left. + (1 - \lambda(D_t, r_t; \theta) h) \cdot V^{\mathsf{B}} \left(t+h, \widehat{\mathbf{U}}_t \cup (r_t, D_t) \right) \Big\} + o(h) \right\},
\end{aligned}
$$

where the second equality enumerates the outcomes during a small time period $[t, t+h]$. Here, $\mathbb{P}(\cdot \mid \widehat{\mathbf{U}}_t)$ is the posterior probability distribution over the possible parameter values in set Θ. The posterior distribution is computed using Bayes' theorem from a prior distribution $\mathbb{P}_0(\theta), \theta \in \Theta$, and the likelihood function $\mathbb{P}(\widehat{\mathbf{U}}_t \mid \theta)$. We note from (A.2) that the uncertainty about the demand parameters is reflected in the posterior distribution that changes over time as more evidence is collected. The uncertainty about the inter-arrival times is modeled by using (II.2.5) to enumerate the probability of zero, one, or more arrivals during the time period $[t, t+h]$.

Bayesian inference is commonly used for deriving the optimal pricing and learning policy in Operations Management literature (see section 4.1.2 in *den Boer* 2015a for a survey of works using a Bayesian approach). This approach is attractive since, as can be seen in (A.2), the chosen pricing decision considers the seller's confidence about the different model alternatives. However, there are several drawbacks to this approach. The main criticism is that the prior distribution is subjective information, and different prior distributions may result in different posteriors and different pricing policies. A second criticism is that, past literature often assumes the set Θ has only two parameter vector values. This assumption aids in deriving an expression for the optimal pricing and learning policy. However, in practice, it is difficult to impose this assumption in practice where parameter values can take on an infinite possibility of values.

A frequentist school of statistical inference does not have these drawbacks. The frequentist school is rooted in the philosophy that there is only one true value for the parameters, so it does not make sense to assign probabilities (prior or posterior) for the different parameter values. For example, the inference from maximum likelihood estimation (MLE) is a point estimate since it chooses the model parameters which has the highest likelihood of observing the data.

If the seller uses a MLE approach, let $V^{\mathsf{MLE}}(t, \widehat{\mathbf{U}}_t)$ be her inferred optimal expected revenue-to-go at time t given data $\widehat{\mathbf{U}_t}$. If $\hat{\theta}(\widehat{\mathbf{U}}_t)$ denotes the ML estimator, then by enumerating outcomes, we can write V^{MLE} as

$$\begin{aligned} V^{\mathsf{MLE}}(t, \widehat{\mathbf{U}}_t) &= \max_{r_t} \left\{ \mathbb{E}_{\hat{\theta}(\widehat{\mathbf{U}}_t)} \left[r_t \cdot (D_{t+h} - D_t) + V^{\mathsf{MLE}} \left(t+h, \widehat{\mathbf{U}}_t \cup (r_t, D_{t+h}) \right) \;\middle|\; D_t \right] \right\} \\ &= \max_{r_t} \left\{ \lambda \left(D_t, r_t; \hat{\theta}(\widehat{\mathbf{U}}_t) \right) h \cdot \left[r_t + V^{\mathsf{MLE}} \left(t+h, \widehat{\mathbf{U}}_t \cup (r_t, D_t + 1) \right) \right] \right. \\ &\qquad \left. + \left(1 - \lambda \left(D_t, r_t; \hat{\theta}(\widehat{\mathbf{U}}_t) \right) h \right) \cdot V^{\mathsf{MLE}} \left(t+h, \widehat{\mathbf{U}}_t \cup (r_t, D_t) \right) + o(h) \right\} \end{aligned} \tag{A.3}$$

We observe from comparing (A.3) to (A.2) that the MLE pricing-and-learning problem infers a point estimate from the data. In contrast, the Bayesian pricing-and-learning problem infers probabilities to the possible parameter values. This contrast is due to the different philosophies of Bayesian inference and frequentist inference. Nevertheless, the similarities between (A.2) and (A.3) are noteworthy. First, both models use an inference about the demand model parameters (informed by the data) to compute the seller's inferred expected revenue-to-go. Specifically, in computing the value-to-go, they do not rely on θ_0 which is unknown to the seller. Second, both models capture parameter uncertainty since they both assume that the data, hence the inference about the parameters, changes over time. Third, they model how prices affect both the current period revenue as well as the data used for future inference.

A.3.1 Optimal pricing and learning policy with ML inference

We will next derive the optimal pricing-and-learning policy that solves (A.3). For ease of exposition, we present our analysis for the case where p_0 and q_0 are unknown, but m_0 is known. Our analysis can be extended to the case where m_0 is also unknown, which we discuss at the end of this subsection.

Remark A.1. If m_0 is known, the log-likelihood function $\mathcal{L}_t$ in (II.4.3) is strongly concave in (p, q) when $\widehat{D}_t \geq 2$. The proof is similar to that of Proposition II.2. We can also check that if $\widehat{D}_t = 0$ or if $\widehat{D}_t = 1$, then $\nabla_{(p,q)}\mathcal{L}_t = 0$ has no solution, so the seller cannot form an inference through MLE. Hence, we will assume in this section that while $D_t < 2$, the seller offers the Markovian Bass prices under some initial guess $\hat{\theta}_0$ of the parameters.

The ML estimator $\hat{\theta}(\widehat{\mathbf{U}}_t) = (\hat{p}_t, \hat{q}_t)$ is the seller's inference about the demand parameters based on the data $\widehat{\mathbf{U}}_t$ at time t. Hence, to maximize the seller's inferred revenue-to-go (A.3), we need to analyze how the change in the state from $(t, \widehat{\mathbf{U}}_t)$ to $(t+h, \widehat{\mathbf{U}}_{t+h})$ in a small time period of length $o(h)$ would affect her future inference $\hat{\theta}(\widehat{\mathbf{U}}_{t+h}) = (\hat{p}_{t+h}, \hat{q}_{t+h})$. This analysis will allow us to reduce the dimension of the state space, resulting in a more tractable decision problem.

To aid in our analysis we assume that the seller forms her inference by solveing the unconstrained model $\max_\theta \mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \theta)$. Hence, the MLE solution satisfies the first-order condition $\nabla_\theta \mathcal{L}_t = 0$. This allows us to prove the following proposition which establishes that the change in the ML estimate $(\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t)$ depends only on the price r^π_{t-}, on the current ML estimate $(\hat{p}_{t-}, \hat{q}_{t-})$, and on the cumulative number of adoptions D^π_{t-}; it does not depend on the entire sample path of adoptions.

Proposition A.1. If $t_2 = \inf\{t : D^\pi_t \geq 2\}$, then $\{(\hat{p}_t, \hat{q}_t),\ t \geq t_2\}$ satisfies the stochastic differential equations:

$$\mathrm{d}\hat{p}_t = \frac{\sigma_p^2 + \sigma_{pq}^2 \frac{D^\pi_{t-}}{m_0}}{\hat{p}_{t-} + \frac{D^\pi_{t-}}{m_0}\hat{q}_{t-}} \left[\mathrm{d}D^\pi_t - \left(\hat{p}_{t-} + \frac{D^\pi_{t-}}{m_0}\hat{q}_{t-} \right) (m_0 - D^\pi_{t-}) x(r^\pi_{t-}) \mathrm{d}t \right], \tag{A.4}$$

$$\mathrm{d}\hat{q}_t = \frac{\sigma_{pq}^2 + \sigma_q^2 \frac{D^\pi_{t-}}{m_0}}{\hat{p}_{t-} + \frac{D^\pi_{t-}}{m_0}\hat{q}_{t-}} \left[\mathrm{d}D^\pi_t - \left(\hat{p}_{t-} + \frac{D^\pi_{t-}}{m_0}\hat{q}_{t-} \right) (m_0 - D^\pi_{t-}) x(r^\pi_{t-}) \mathrm{d}t \right]. \tag{A.5}$$

Here, variables $\sigma_p^2, \sigma_q^2, \sigma_{pq}^2$ are defined as follows:

$$\begin{pmatrix} \sigma_p^2 & \sigma_{pq}^2 \\ \sigma_{pq}^2 & \sigma_q^2 \end{pmatrix} := \begin{pmatrix} \dfrac{1}{1-\cos^2\alpha} \cdot \dfrac{1}{\|\mathbf{v}_p\|^2} & -\dfrac{\cos^2\alpha}{1-\cos^2\alpha} \cdot \dfrac{1}{\mathbf{v}_p^\top \mathbf{v}_q} \\ -\dfrac{\cos^2\alpha}{1-\cos^2\alpha} \cdot \dfrac{1}{\mathbf{v}_p^\top \mathbf{v}_q} & \dfrac{1}{1-\cos^2\alpha} \cdot \dfrac{1}{\|\mathbf{v}_q\|^2} \end{pmatrix}, \tag{A.6}$$

where $\mathbf{v}_p$ and $\mathbf{v}_q$ are D^π_{t-}-dimensional column vectors with entries

$$v_{p,j} = \frac{1}{\hat{p}_{t-} + \frac{j-1}{m_0}\hat{q}_{t-}}, \quad v_{q,j} = \frac{\frac{j-1}{m_0}}{\hat{p}_{t-} + \frac{j-1}{m_0}\hat{q}_{t-}}, \quad \text{for } j = 1, 2, \ldots, D^\pi_{t-}$$

and α is the angle between vectors $\mathbf{v}_p$ and $\mathbf{v}_q$.

Note that $\sigma_p^2, \sigma_q^2, \sigma_{pq}^2$ in (A.6) can be interpreted as *estimated* variances and covariances of estimated parameters (the actual variance-covariance matrix contains the true values of p_0 and q_0). Note that since $\|\mathbf{v}_p\|^2$ and $\|\mathbf{v}_q\|^2$ are in the order of $\Theta\left(\frac{1}{D_t^\pi+1}\right)$ and $0 < \alpha < \frac{\pi}{2}$, so are $\sigma_p^2, \sigma_q^2, \sigma_{pq}^2$. Therefore, the updates $\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t$ to the ML estimators will approach zero as adoptions increase. This observation is consistent with Lemma II.3.

We prove Proposition A.1 by noting that the value of the log-likelihood function $\mathcal{L}_t$ is continuously changing in time, with jumps for each new adoption. Hence, the differential $(\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t)$ is impacted by a continuous phenomenon and a discrete phenomenon. We apply Itô's lemma to derive the stochastic differential equation of $(\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t)$. Intuitively speaking, Itô's lemma is a chain rule defined on stochastic processes. Since the first-order condition $\nabla_\theta \mathcal{L}_t|_{\theta=\hat{\theta}_t} = 0$ is an implicit function of $\hat{\theta}_t = (\hat{p}_t, \hat{q}_t)$, we apply Itô's lemma on this implicit function. Note that in the proposition, the differential equation holds for $t \geq t_2$, which ensures that the first-order condition can be met by a unique solution (see Remark A.1).

The literature on learning under continuous-time stochastic processes typically focuses on deriving asymptotic properties (such as weak consistency and asymptotic normality) of ML estimators in a CTMC. *Zhao and Xie* (1996) study MLE for nonhomogeneous Poisson processes where the transition rates are only affected by time. *Keiding et al.* (1975) and *Küchler and Sorensen* (2006) study estimations for simple linear birth-death processes. In contrast, Proposition A.1 establishes the evolution over time of ML estimators to the Markovian Bass model. Note that the martingale property of $(\hat{p}_t, \hat{q}_t)$ is consistent with the general results of the partial derivatives of log-likelihood functions for exponential families derived in *Küchler and Sorensen* (2006) Theorem 8.1.2.

An implication of Proposition A.1 is that the state variables $(D_{t-}^\pi, t, \hat{p}_{t-}, \hat{q}_{t-})$ encapsulate the information sufficient to choose a price that maximizes the seller's inferred revenue-to-go function in (A.3). Given $(D_{t-}^\pi, \hat{p}_{t-}, \hat{q}_{t-})$, Proposition A.1 states that $\mathrm{d}\hat{p}_t, \mathrm{d}\hat{q}_t$ are random variables that depend on $\mathrm{d}D_t^\pi$. But given state $(t, \widehat{\mathbf{U}}_t)$ in (A.3), the seller infers $\mathrm{d}D_t^\pi$ to follow a Markovian Bass model with parameter $\hat{\theta}_{t-}$. Hence, the seller's inferred revenue-to-go function (A.3) only depends on the state through variables $(D_{t-}^\pi, t, \hat{p}_{t-}, \hat{q}_{t-})$.

Hence, we can reduce the dimension of V^{MLE} in (A.3). Let $V^{\mathsf{MLE}}(d, t, p, q)$ denote the seller's inferred revenue-to-go where d is the cumulative number of adoptions, t is the time remaining in the horizon, and (p, q) are the current ML estimators. For any $d \geq 2$, we have

$$
\begin{aligned}
&V^{\mathsf{MLE}}(d,t,p,q) \\
&= \max_{r} \quad \Big\{ \Big(r + V^{\mathsf{MLE}}(d+1, t-\delta t, p+\eta_p+\mu_p(r)\delta t, q+\eta_q+\mu_q(r)\delta t)\Big) \cdot \lambda(d,r;p,q)\delta t \\
&\quad + V^{\mathsf{MLE}}(d, t-\delta t, p+\mu_p(r)\delta t, q+\mu_q(r)\delta t) \cdot [1-\lambda(d,r;p,q)\delta t] + o(\delta t)\Big\},
\end{aligned} \tag{A.7}
$$

where

$$
\mu_p(r) := -\left(\sigma_p^2 + \sigma_{pq}^2 \frac{d}{m_0}\right)(m_0-d)x(r), \qquad \eta_p := \frac{\sigma_p^2 + \sigma_{pq}^2 \frac{d}{m_0}}{p+\frac{d}{m_0}q}, \tag{A.8}
$$

$$
\mu_q(r) := -\left(\sigma_{pq}^2 + \sigma_q^2 \frac{d}{m_0}\right)(m_0-d)x(r), \qquad \eta_q := \frac{\sigma_{pq}^2 + \sigma_q^2 \frac{d}{m_0}}{p+\frac{d}{m_0}q}, \tag{A.9}
$$

and $\sigma_p^2, \sigma_q^2, \sigma_{pq}^2$ are defined in (A.6) with $\hat{p}_{t-} = p$, $\hat{q}_{t-} = q$, and $D_{t-}^{\pi} = d$.

The next result characterizes the optimal pricing-and-learning policy, which we refer as the MBP-MLE-Learning policy.

Theorem A.2 (MBP-MLE-Learning)**.** Let $r^*(d,t,p,q)$ be the price offered under the optimal policy π^* to the Markovian Bass pricing-and-learning problem (A.1) when $d \in \{2, 3, \ldots, m_0 - 1\}$ is the total past sales, $t \in [0, T]$ is the time remaining in the sales horizon, and the seller uses MLE to infer demand with (p, q) as the current ML estimators. Then $r^*(d,t,p,q)$ is the solution to the equation

$$
r = -\frac{x(r)}{x'(r)} - V^{\mathsf{MLE}}(d+1, t, p+\eta_p, q+\eta_q) + V^{\mathsf{MLE}}(d,t,p,q) + \frac{\partial V^{\mathsf{MLE}}}{\partial p}\eta_p + \frac{\partial V^{\mathsf{MLE}}}{\partial q}\eta_q, \tag{A.10}
$$

where η_p,η_q are defined in (A.8),(A.9), and $V^{\mathsf{MLE}}(\cdot,\cdot,\cdot,\cdot)$ is a function that solves the HJB differential equation

$$
\frac{\partial}{\partial t} V^{\mathsf{MLE}}(d,t,p,q) + \frac{x(r^*(d,t,p,q))^2}{x'(r^*(d,t,p,q))}\left(p + \frac{d}{m_0}q\right)(m_0 - d) = 0, \tag{A.11}
$$

with boundary conditions $V^{\mathsf{MLE}}(m_0,t,p,q) = 0$ for all $t \in [0,T]$, $(p,q) \in \Re^2$, and $V^{\mathsf{MLE}}(d,0,p,q) = 0$ for all $d \in \{0, 1, \ldots, m_0\}$, $(p,q,) \in \Re^2$.

Note that (A.10) could be rearranged as:

$$r^* = \underbrace{-\frac{x(r^*)}{x'(r^*)} - V^{\mathsf{MLE}}(d+1,t,p,q) + V^{\mathsf{MLE}}(d,t,p,q)}_{\text{Price exploitation}}$$
$$+ \underbrace{V^{\mathsf{MLE}}(d+1,t,p,q) - V^{\mathsf{MLE}}(d+1,t,p+\eta_p,q+\eta_q) + \frac{\partial V^{\mathsf{MLE}}}{\partial p}\eta_p + \frac{\partial V^{\mathsf{MLE}}}{\partial q}\eta_q}_{\text{Price exploration } \left(=0 \text{ when } \sigma_p^2,\sigma_q^2,\sigma_{pq}^2=0\right)} . \tag{A.12}$$

Hence, the MBP-MLE-Learning policy has two components. The first is the Markovian Bass price (MBP) in (II.3.2), but using the current parameter estimates. This component maximizes the revenue by exploiting the current information. Because of estimation errors, the firm experiments with prices that allow it to learn quickly. This is captured in the second component, which acts as a "perturbation" to allow the firm to improve estimation by setting a price possibly lower than MBP. Note that the perturbation term fades over time as the estimates become more accurate.

The values for V^{MLE} can be found numerically using a method such as finite differences (see Appendix A.1).

Remark A.2. The pricing-and-learning setting is related to restless bandits problems where the rewards of all arms evolve (*Whittle*, 1988; *Cohen and Solan*, 2013). This resembles the evolving rate of adoptions over time because of changing cumulative adoptions, which affects the revenue rate of all possible prices (arms). In addition to differences in the problem setup (continuous time and price), a key difference in our setting is that learning does not occur at the same rate: some prices (arms) induce higher adoption rates and accelerate learning (see Lemma II.3). In the bandit problems, this is equivalent to the case where some arms let you learn faster.

Remark A.3. Thus far, we assumed that p_0 and q_0 are unknown, but the value of m_0 is known. If m_0 is also unknown, Proposition A.1 can be extended by transforming $\hat{\theta}_t = (\hat{p}_t, \hat{q}_t, \hat{m}_t)$ into $\hat{\beta}_t = (\hat{\beta}_{t1}, \hat{\beta}_{t2}, \hat{\beta}_{t3})$ according to (II.4.4). Deriving the stochastic differential equation for $\hat{\beta}_t$ follows exactly the same procedure as the proof of Proposition A.1. This is because the adoption rate is linear in $(\beta_1, \beta_2, \beta_3)$, which is analogous to the case where the adoption rate is linear in (p, q) if m_0 is known. Accordingly, we can derive an analogue to Theorem A.2 which is the optimal pricing-and-learning policy when the parameters $(\beta_{00}, \beta_{01}, \beta_{02})$ are inferred from maximum likelihood.

APPENDIX B

Proofs of chapter III

B.1 Companion results

Lemma B.1. Define $r(y) := x^{-1}(y)y$. Under Assumption III.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.)

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 III.2(ii) the denominator is negative. Assumption III.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 III.2(iv),(vii). $\square$

B.2 Section III.3 proofs

B.2.1 Proof of Theorem III.1

Proof. Any feasible solution to (**D**) is also feasible in ($\mathbf{D}^\dagger$), so $V^{\mathsf{D}}(T;u,\alpha) \leq V^{\mathsf{D}^\dagger}(T;u,\alpha)$. To show "$\geq$", 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 (**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 III.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. □

B.2.2 Proof of Theorem III.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

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

To proceed, we require the following claim.

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

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

We now show that each term in the summation of (B.1) is jointly concave in $(d_1, d_2, \ldots, d_T)$. Consider any $\theta \in [0,1]$, $d^1 = (d_1^1, d_2^1, \ldots, d_T^1)$ and $d^2 = (d_1^2, d_2^2, \ldots, d_T^2)$. We define the vector $\bar{d} \triangleq \theta d^1 + (1-\theta) d^2$, where $\bar{d}_t = \theta d_t^1 + (1-\theta) d_t^2$.
Consider an arbitrary index t. Because $\lambda(n, \alpha)$ is jointly concave in (n, α) by Assumption III.2(vi), then

$$\begin{aligned} &\lambda(u - \bar{d}_1 - \bar{d}_2 - \ldots - \bar{d}_{t-1}, \alpha) \\ &\quad \geq \underbrace{\theta \lambda(u - d_1^1 - d_2^1 - \ldots - d_{t-1}^1, \alpha) + (1-\theta)\lambda(u - d_1^2 - d_2^2 - \ldots - d_{t-1}^2, \alpha)}_{\bar{\lambda}}. \end{aligned} \tag{B.2}$$

From the definition of r, we have that

$$\begin{aligned} &\lambda(u - \bar{d}_1 - \bar{d}_2 - \ldots - \bar{d}_{t-1}, \alpha) \cdot r\left(\frac{\bar{d}_t}{\lambda(u - \bar{d}_1 - \bar{d}_2 - \ldots - \bar{d}_{t-1}, \alpha)}\right) \\ &= \bar{d}_t \cdot x^{-1}\left(\frac{\bar{d}_t}{\lambda(u - \bar{d}_1 - \ldots - \bar{d}_{t-1}, \alpha)}\right) \\ &\geq \bar{d}_t \cdot x^{-1}\left(\frac{\bar{d}_t}{\bar{\lambda}}\right) = \bar{\lambda} \cdot r\left(\frac{\bar{d}_t}{\bar{\lambda}}\right) \\ &> \theta \lambda(u - d_1^1 - \ldots - d_{t-1}^1, \alpha) \cdot r\left(\frac{d_t^1}{\lambda(u - d_1^1 - \ldots - d_{t-1}^1, \alpha)}\right) \\ &\quad + (1-\theta)\lambda(u - d_1^2 - \ldots - d_{t-1}^2, \alpha) \cdot r\left(\frac{d_t^2}{\lambda(u - d_1^2 - \ldots - d_{t-1}^2, \alpha)}\right) \end{aligned}$$

where the first inequality follows from the fact that x^{-1} is a monotone decreasing function and from (B.2). The second inequality follows Claim B.1. Hence, this shows that each term in the summation (B.1) 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 $\bar{d}_t \leq \lambda(u - \bar{d}_1 - \ldots - \bar{d}_{t-1}, \alpha)$ for all t.

$$\bar{d}_t = \theta d_t^1 + (1-\theta) d_t^2$$

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

where the first inequality follows from the feasibility of d^1 and d^2, and the second inequality follows from (B.2). 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:

$$
\begin{aligned}
V^{\mathsf{D}}(T; u, \alpha) = \max_{n,d} \quad & \sum_{t=1}^{T} x^{-1}\left(\frac{d_t}{\lambda(n_{t-1}, \alpha)}\right) \cdot d_t \\
\text{s.t.} \quad & \sum_{t=1}^{T} d_t \leq u \\
& n_t = n_{t-1} - d_t \quad \text{for all } t \geq 1 \\
& n_0 = u \\
& 0 \leq d_t \leq \lambda(n_{t-1}, \alpha) \quad \text{for all } t \geq 1.
\end{aligned}
\tag{B.3}
$$

For any $(u_1, \alpha_1) \geq 0$ and $(u_2, \alpha_2) \geq 0$, 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 $(n^1, d^1) \neq (n^2, d^2)$. 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 (B.3) 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 (B.3) with $u = \bar{u}$ and $\alpha = \bar{\alpha}$. 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

$$\begin{aligned}\sum_{t=1}^{T} g(\bar{d}_t, \bar{n}_t) &> \theta \cdot \sum_{t=1}^{T} x^{-1}\left(\frac{d_t^1}{\lambda(n_t^1, \alpha_1)}\right) d_t^1 + (1-\theta) \cdot \sum_{t=1}^{T} x^{-1}\left(\frac{d_t^2}{\lambda(n_t^2, \alpha_2)}\right) d_t^2 \\ &= \theta V^{\mathsf{D}}(\alpha_1, T) + (1-\theta) V^{\mathsf{D}}(\alpha_2, T).\end{aligned} \tag{B.4}$$

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

$$g(\bar{d}_t, \bar{n}_t) > \theta x^{-1}\left(\frac{d_t^1}{\lambda(n_t^1, \alpha_1)}\right) d_t^1 + (1-\theta) x^{-1}\left(\frac{d_t^2}{\lambda(n_t^2, \alpha_2)}\right) d_t^2. \tag{B.5}$$

To show (B.5), 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 B.1.

We now show (B.5), because $\lambda(n, \alpha)$ is jointly concave in (n, α) by Assumption III.2(vi), hence $\lambda(\bar{n}_t, \bar{\alpha}) \geq \theta\lambda(n_t^1, \alpha_1) + (1-\theta)\lambda(n_t^2, \alpha_2)$. Then because x^{-1} is a monotone decreasing function, we have

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

where the first equality is from the definition of r, and the last inequality is from Claim B.1. This establishes (B.5), which in turn yields (B.4). This completes the proof.

□

B.2.3 Proof of Theorem III.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 III.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**):

$$R^{\mathsf{D}}(u,T) = \max_{y} \quad x^{-1}(y)\lambda(u,\alpha)y + R^{\mathsf{D}}\left(u - \lambda(u,\alpha)y, T-1\right) \tag{B.6a}$$

$$\text{s.t.} \quad \lambda(u,\alpha)y \le u, \tag{B.6b}$$

where the base case is $R^{\mathsf{D}}(u,0)$ 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 (B.6). Hence, to show that (**D**) has a unique solution, we need to show that (B.6) has a unique solution. Since the feasible set of (B.6) is compact, to show that (B.6) has a unique solution, it suffices to show that the objective function,

$$R^{\mathsf{D},y}(u,T) \triangleq x^{-1}(y)\lambda(u,\alpha)y + R^{\mathsf{D}}\left(u - \lambda(u,\alpha)y, T-1\right) \tag{B.7}$$

is strictly concave in y.

Claim B.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 B.1(i). To see that the second term is also concave, its second-order derivative with respect to y is

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

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

□

B.2.4 Proof of Theorem III.4

The proof requires the following lemma.

Lemma B.2. 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 B.2. 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 B.1

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

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 $t \ge i+1$ 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 B.1 (MOVE(i)) constructs y'. ☐

Now we can proceed with the proof of the theorem.

Proof of Theorem III.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 first 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 $x^{-1}\left(u/\lambda(u, \alpha)\right)u > 0$. 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 B.2, we can construct a different solution with the same objective value. This contradicts Theorem III.3 that the optimal solution of (**D**) is unique. □

B.2.5 Strong duality of dynamic programming counterpart of (D)

For a fixed α, note that $R^{\mathsf{D}}(u, T)$ in (B.6) 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 III.1.

Lemma B.3. Fix α. For any $u \in (0, \alpha]$,

$$R^{\mathsf{D}}(u, T) = \inf_{\mu \geq 0} L^{\mathsf{D},\mu}(u, T), \tag{B.8}$$

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

$$L^{\mathsf{D},\mu}(u, T) \triangleq \max_{y \in [0,1]} \left\{ x^{-1}(y)\lambda(u, \alpha)y + R^{\mathsf{D}}\left(u - \lambda(u, \alpha)y,\ T - 1\right) + \mu\left(u - \lambda(u, \alpha)y\right) \right\}. \tag{B.9}$$

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 (B.6) is a convex optimization problem with a feasible point that satisfies its constraints strictly. Observe that all the constraints in (B.6) are affine in y. The objective function is concave in y, as established in Claim B.2. Hence, (B.6) is a convex optimization problem

The next step is to demonstrate that there exists a feasible solution to (B.6) that satisfies the inequality constraint (B.6b) strictly. Notice that any $y \in (0, \min\{1, u/\lambda(u, \alpha)\})$ is strictly feasible to (B.6) because since $u > 0$ and with Assumption III.2(v), $u/\lambda(u, \alpha) > 0$. Hence, Slater's condition implies (B.8) holds. □

B.2.6 Proof of Proposition III.1

Proof. We first introduce the dynamic programming counterpart of ($\mathbf{D}^{\dagger}$) for any $u \in [0, \alpha]$:

$$R^{\mathsf{D}_0}(u, T) \quad \triangleq \quad \max_{y \in [0,1]} \quad x^{-1}(y) \min\left(\lambda(u, \alpha)y, u\right) + R^{D_0}\left(\left[u - \lambda(u, \alpha)y\right]^+,\ T - 1\right).$$

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:

$$R^*(u,1) \triangleq \max_{y\in[0,1]} \mathbb{E}_{y,u}\left[x^{-1}(y)\min(D,u)\right] \tag{B.10a}$$

For a given y, let us denote the objective value (B.10a) as $V^y(u,1)$.

Consider any $y \in [0,1]$. We have that

$$\begin{aligned}
V^y(u,1) &= \mathbb{E}_{y,u}\left[x^{-1}(y)\min(D,u)\right] \\
&\leq \max_{y_0\in[0,1]} \mathbb{E}_{y_0,u}\left[x^{-1}(y_0)\min(D,u)\right] \\
&= \max_{y_0\in[0,1]} x^{-1}(y_0)\mathbb{E}_{y_0,u}\left[\min(D,u)\right] \\
&\leq \max_{y_0\in[0,1]} x^{-1}(y_0)\min\left(\mathbb{E}_{y_0,u}(D),u\right) &\text{(B.11)} \\
&= \max_{y_0\in[0,1]} x^{-1}(y_0)\min\left(\lambda(u,\alpha)y_0,u\right). &\text{(B.12)}
\end{aligned}$$

Here, (B.11) comes from $\min(D,n)$ is a concave function of D and Jensen's inequality. (B.12) 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 (B.11) is equal to $R^{\mathsf{D}_0}(u,1)$. Therefore, we have that

$$V^y(u,1) \leq R^{\mathsf{D}_0}(u,1). \tag{B.13}$$

The last step to finish the base case of induction is to take the supremum of the left-hand side of (B.13) over all $y \in [0,1]$. This yields $R^*(u,1) \leq R^{\mathsf{D}_0}(u,1) = R^{\mathsf{D}}(u,1)$.

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:

$$\begin{aligned}
R^*(u,T'+1) = \max_{y\in[0,1]} \ & \mathbb{E}_{y,u}\left[x^{-1}(y)\left(D-[D-u]^+\right) + R^*([u-D]^+,T')\right] &\text{(B.14a)} \\
\text{s.t. } & \mathbb{E}_{y,u}(D) = \lambda(u,\alpha)y. &\text{(B.14b)}
\end{aligned}$$

Claim B.3. The maximization problem (B.14) is feasible and $R^*(u,T'+1)$ is bounded.

We know $y=0$ is a feasible solution. Moreover, the objective function (B.14a) 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 B.1(ii). This concludes the claim.

Now, consider any $y \in [0,1]$ feasible to (B.14b). We denote its objective value (B.14a) as $V^y(u, T'+1)$. Then for any γ, we have that

$$V^y(u,T'+1) \leq \mathbb{E}_{y,u}\left[x^{-1}(y)\left(D - [D-u]^+\right) + R^*([u-D]^+, T')\right] + \gamma\left(\mathbb{E}_{y,u}(D) - \lambda(u,\alpha)y\right) \tag{B.15a}$$

$$\leq \max_{y_0\in[0,1]} \mathbb{E}_{y_0,u}\Big[x^{-1}(y_0)\left(D - [D-u]^+\right) + R^*([u-D]^+, T') + \gamma\left(D - \lambda(u,\alpha)y_0\right)\Big] \tag{B.15b}$$

$$\leq \max_{y_0\in[0,1]} \mathbb{E}_{y_0,u}\Big[x^{-1}(y_0)\left(D - [D-u]^+\right) + R^{\mathsf{D}_0}([u-D]^+, T') + \gamma\left(D - \lambda(u,\alpha)y_0\right)\Big]. \tag{B.15c}$$

Here, (B.15c) comes from the inductive hypothesis. Since (B.15) is true for all feasible y, taking the supremum of $V(y;u,T'+1)$ over $y\in[0,1]$ satisfying (B.14b), we have that $R^*(u,T'+1)$ is bounded above by (B.15c).

Note that (B.15c), and hence $R^*(u,T'+1)$, is bounded above by

$$\max_{\substack{y_0\in[0,1],\\ d\in\Re}} \left\{x^{-1}(y_0)\left(d - [d-u]^+\right) + R^{\mathsf{D}_0}([u-d]^+, T') + \gamma\left(d - \lambda(u,\alpha)y_0\right)\right\}. \tag{B.16}$$

Note that (B.16) is an upper bound because d, being a decision variable that can take any value, results in a larger value than (B.15c). Since (B.16) 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 (B.17) as follows:

$$R^*(u,T'+1) \leq \inf_{\gamma} \max_{\substack{y_0\in[0,1],\\ d\in\Re}} \left\{x^{-1}(y_0)\left(d - [d-u]^+\right) + R^{\mathsf{D}_0}([u-d]^+, T') + \gamma\left(d - \lambda(u,\alpha)y_0\right)\right\}. \tag{B.17}$$

Next, we will prove that the right-hand side of (B.17) equals $R^{\mathsf{D}}(u,T'+1)$. Note that $\gamma = 0$ is the solution to (B.17) because otherwise, d can be chosen such that the value of (B.17) is $+\infty$. Then, for the problem in (B.17), it suffices to restrict $d \leq u$, since any $d > u$ does not improve the value of the objective function. Thus, we know for any $\mu \geq 0$, the right-hand side of (B.17) is upper bounded by

$$\inf_{\gamma} \max_{\substack{y_0\in[0,1],\\ d\leq u}} \left\{x^{-1}(y_0)\left(d - [d-u]^+\right) + R^{\mathsf{D}_0}([u-d]^+, T') + \gamma\left(d - \lambda(u,\alpha)y_0\right) + \mu(u-d)\right\}. \tag{B.18}$$

Because (B.18) is the upper bound of (B.17) for any $\mu \geq 0$, we can take the infimum of (B.18) and yield the final upper bound of (B.17) as follows

$$R^*(u, T'+1) \leq \inf_{\gamma,\mu \geq 0} \max_{\substack{y_0 \in [0,1], \\ d \leq u}} \left\{ x^{-1}(y_0)d + R^{\mathsf{D}_0}(u-d, T') + \gamma\left(d - \lambda(u,\alpha)y_0\right) + \mu(u-d) \right\}. \tag{B.19}$$

Since R^{D} is equivalent to R^{D_0}, we observe that the right-hand side of (B.19) is the dual problem of R^{D} and according to Lemma B.3, we can simplify (B.19) as

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

This finishes our inductive step. □

B.2.7 Proof of Theorem III.5

Proof. It suffices to consider a two-period setting ($T = 2$) and we set $\alpha = \lambda(\alpha,\alpha)y_1^{\mathsf{D}} + \lambda(n_1^{\mathsf{D}},\alpha)y_2^{\mathsf{D}}$ where $y_2^{\mathsf{D}} = \arg\max_{y \in [0,1]} x^{-1}(y)y$ and $\alpha > \lambda(\alpha,\alpha)y_1^{\mathsf{D}} + \sqrt{\lambda(\alpha,\alpha)y_1^{\mathsf{D}}}$. This α is a fixed point such that when demand is deterministic, under the optimal deterministic policy, it uses up all the inventory.

Let $\lambda(\cdot,\cdot)$ be a general homogeneous function with degree 1. Hence, $\lambda(nm,\alpha m) = m\lambda(n,\alpha)$. So, if we let $\lambda^m = \lambda$ for all m, then Assumption III.4 is satisfied.

Suppose that demand follows a three-point distribution such that for any $t = 1, 2$, given y_t and $\mathcal{F}_{t-1}$, the conditional probability of D_t^m is:

$$\begin{aligned}
&\mathbb{P}\left(D_t^m = \lambda(N_{t-1}^m, \alpha m)y_t - \sqrt{\lambda^m(N_{t-1}^m, \alpha m)y_t} \mid \mathcal{F}_{t-1} \right) = 1/3, \\
&\mathbb{P}\left(D_t^m = \lambda(N_{t-1}^m, \alpha m)y_t \mid \mathcal{F}_{t-1} \right) = 1/3, \\
&\mathbb{P}\left(D_t^m = \lambda(N_{t-1}^m, \alpha m)y_t + \sqrt{\lambda(N_{t-1}^m, \alpha m)y_t} \mid \mathcal{F}_{t-1} \right) = 1/3.
\end{aligned}$$

By construction, $\mathbb{E}[D_t^m \mid \mathcal{F}_{t-1}] = \lambda(N_{t-1}^m, \alpha m)y_t$ and this distribution satisfies Assumption III.3 because $\mathrm{Var}(D_t^m | \mathcal{F}_{t-1}) = \frac{2}{3}\lambda^m(N_{t-1}^m, \alpha m)y_t$.

We prove the theorem by analyzing the expected revenue loss under each outcome of D_1^m. Note that, by definition of policy $\mathsf{CE\text{-}CL}$, we have $y_1^{\mathsf{CL}} = y_1^{\mathsf{D}}$. Also, $N_0^m = \alpha m$.

Let $\mathcal{E}_0$ denote the event $\{D_1^m = \lambda(N_0^m, \alpha m)y_1^{\mathsf{CL}}\} = \{D_1^m = \lambda(\alpha m, \alpha m)y_1^{\mathsf{D}}\}$. Under $\mathcal{E}_0$, there is no stockout in period 1 due to our choice of α, so the period 1 revenue of $\mathsf{CE\text{-}CL}$ is equal to $x^{-1}(y_1^{\mathsf{CL}})D_1^m = x^{-1}(y_1^{\mathsf{D}})\lambda(\alpha m, \alpha m)y_1^{\mathsf{D}}$. Note that this coincides with the period 1 revenue in $V^{\mathsf{D}}(m, T)$. So given event $\mathcal{E}_0$, the period 1 revenue loss of $\mathsf{CE\text{-}CL}$ is zero.

Under $\mathcal{E}_0$, since there is no period 1 stockout, then together with the constraints of (**D**) and the fact that λ is homogeneous with degree 1, we have:

$$N_1^m = \alpha m - D_1^m = \alpha m - \lambda(\alpha m, \alpha m) y_1^{\mathsf{D}} = m\left(\alpha - \lambda(\alpha, \alpha) y_1^{\mathsf{D}}\right) = m n_1^{\mathsf{D}}. \tag{B.20}$$

Recall that $n_1^{\mathsf{D},m} = m n_1^{\mathsf{D}}$. So, under $\mathsf{CE\text{-}CL}$, the remaining inventory at the end of period 1 is the same as that under the deterministic model $V^{\mathsf{D}}(m,T)$. This implies that, under event $\mathcal{E}_0$, we have that $y_2^{\mathsf{CL}} = y_2^{\mathsf{D}}$.

Combining these observations, we know that under $\mathcal{E}_0$, the conditional expected revenue loss of $\mathsf{CE\text{-}CL}$ is equal to:

$$\begin{aligned}&0 + \mathbb{E}\left[x^{-1}(y_2^{\mathsf{D}})\lambda(m n_1^{\mathsf{D}}, \alpha m) y_2^{\mathsf{D}} - x^{-1}(y_2^{\mathsf{CL}}) \min(D_2^m, N_1^m) \mid \mathcal{E}_0\right] \\ &= x^{-1}(y_2^{\mathsf{D}})\lambda(m n_1^{\mathsf{D}}, \alpha m) y_2^{\mathsf{D}} - x^{-1}(y_2^{\mathsf{D}})\mathbb{E}\left[N_1^m - [N_1^m - D_2^m]^+ \mid \mathcal{E}_0\right].\end{aligned} \tag{B.21}$$

From (B.20), $N_1^m = m(\alpha - \lambda(\alpha,\alpha) y_1^{\mathsf{D}})$ on $\mathcal{E}_0$, which implies from our choice of α that $N_1^m = m\lambda(n_1^{\mathsf{D}}, \alpha) y_2^{\mathsf{D}} = \lambda(m n_1^{\mathsf{D}}, \alpha m) y_2^{\mathsf{D}}$. Thus, (B.21) reduces to

$$\begin{aligned}&x^{-1}(y_2^{\mathsf{D}})\lambda(m n_1^{\mathsf{D}}, \alpha m) y_2^{\mathsf{D}} - x^{-1}(y_2^{\mathsf{D}})\lambda(m n_1^{\mathsf{D}}, \alpha m) y_2^{\mathsf{D}} + x^{-1}(y_2^{\mathsf{D}})\mathbb{E}\left([N_1^m - D_2^m]^+ \mid \mathcal{E}_0\right) \\ &= x^{-1}(y_2^{\mathsf{D}})\mathbb{E}\left([\lambda(m n_1^{\mathsf{D}}, \alpha m) y_2^{\mathsf{D}} - D_2^m]^+ \mid \mathcal{E}_0\right) \\ &= x^{-1}(y_2^{\mathsf{D}})\mathbb{E}\left([\lambda(N_1^m, \alpha m) y_2^{\mathsf{D}} - D_2^m]^+ \mid \mathcal{E}_0\right) \\ &= \frac{1}{3} x^{-1}(y_2^{\mathsf{D}})\sqrt{\lambda(m n_1^{\mathsf{D}}, \alpha m) y_2^{\mathsf{D}}} = \Theta(\sqrt{m}).\end{aligned} \tag{B.22}$$

Hence, given $\mathcal{E}_0$, the conditional expected revenue loss of $\mathsf{CE\text{-}CL}$ is $\Theta(\sqrt{m})$.

Denote the two events:

$$\left\{D_1^m = \lambda(\alpha m, \alpha m) y_1^{\mathsf{CL}} - \sqrt{\lambda(\alpha m, \alpha m) y_1^{\mathsf{CL}}}\right\} \text{ and } \left\{D_1^m = \lambda(\alpha m, \alpha m) y_1^{\mathsf{CL}} + \sqrt{\lambda(\alpha m, \alpha m) y_1^{\mathsf{CL}}}\right\}$$

as $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively. Let r_1 and r_2 denote the conditional expected revenue of $\mathsf{CE\text{-}CL}$ given $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively. Here, the conditional expectation is with respect to the three-point demand process.

Consider a new demand process. For $t = 1$, demand follows a two-point distribution such that given y_t, the conditional probability of D_1^m is:

$$\begin{aligned}\mathbb{P}\left(D_1^m = \lambda(N_0^m, \alpha m) y_1 - \sqrt{\lambda(N_0^m, \alpha m) y_1} \mid \mathcal{F}_0\right) = 1/2, \\ \mathbb{P}\left(D_1^m = \lambda(N_0^m, \alpha m) y_1 + \sqrt{\lambda(N_0^m, \alpha m) y_1} \mid \mathcal{F}_0\right) = 1/2.\end{aligned}$$

For $t = 2$, demand follows the three-point distribution introduced earlier.

Note that the difference between the old and the new demand processes is only the demand distribution in period 1. So, given the period 1 demand realization, the conditional expected revenue of CE-CL is the same under both processes. Hence, r_1 (r_2) is also the conditional expected revenue of CE-CL given $\mathcal{E}_1$ ($\mathcal{E}_2$) under the new demand process. Since now $\Omega = \mathcal{E}_1 \cup \mathcal{E}_2$, then $r_1/2 + r_2/2$ is the expected revenue of CE-CL under the new process.

Hence, if r^* is the optimal expected revenue under the new demand process, then:

$$\frac{1}{3}r_1 + \frac{1}{3}r_2 = \frac{2}{3}\left(\frac{r_1}{2} + \frac{r_2}{2}\right) \leq \frac{2}{3}r^* \leq \frac{2}{3}V^{\mathsf{D}}(m,T). \tag{B.23}$$

Here, the last inequality comes from Proposition III.1 and from the fact that $V^{\mathsf{D}}(m,T)$ is also the deterministic model under the new process.

Therefore, putting (B.22) and (B.23) together, we have the expected revenue loss of CE-CL satisfies

$$\begin{aligned}
&V^{\mathsf{D}}(m,T) - V^{\mathsf{CL}}(m,T) \\
&= \frac{1}{3}x^{-1}(y_2^{\mathsf{D}})\sqrt{\lambda(mn_1^{\mathsf{D}}, \alpha m)y_2^{\mathsf{D}}} + \frac{1}{3}(V^{\mathsf{D}}(m,T) - r_1) + \frac{1}{3}(V^{\mathsf{D}}(m,T) - r_2) \\
&\geq \frac{1}{3}x^{-1}(y_2^{\mathsf{D}})\sqrt{\lambda(mn_1^{\mathsf{D}}, \alpha m)y_2^{\mathsf{D}}} + 0 = \Theta(\sqrt{m}).
\end{aligned}$$

This completes the proof. □

B.2.8 Proof of Lemma III.1

Proof. When demand is deterministic, Lemma III.1 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, (III.4.6) and (III.4.7) hold for $t = 0$.

For the induction step, assume that (III.4.6) and (III.4.7) hold for $t - 1$, i.e.,

$$\mathbb{E}\left|\bar{I}_{t-1}^m - n_{t-1}^{\mathsf{D}}\right| \leq \Theta(1/\sqrt{m}) \tag{B.24}$$

$$\mathbb{E}\left|\lambda(\bar{I}_{t-1}^m, \alpha) - \lambda(n_{t-1}^{\mathsf{D}}, \alpha)\right| \leq \Theta(1/\sqrt{m}). \tag{B.25}$$

We prove that both properties hold for t.

To prove (III.4.6) for t, notice that by adding and subtracting $\mathbb{E}\left(\bar{I}_t^m\right)$,

$$\mathbb{E}\left|\bar{I}_t^m - n_t^{\mathsf{D}}\right| = \mathbb{E}\left|\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right) + \mathbb{E}\left(\bar{I}_t^m\right) - n_t^{\mathsf{D}}\right| \leq \mathbb{E}\left|\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right| + \left|\mathbb{E}\left(\bar{I}_t^m\right) - n_t^{\mathsf{D}}\right|. \tag{B.26}$$

We will show that both terms in (B.26) are $\mathcal{O}(1/\sqrt{m})$.

Consider first term, $\mathbb{E}\left|\bar{I}_t^m - \mathbb{E}\left(\bar{I}_t^m\right)\right|$. Note that $\mathbb{P}\left(\bar{I}_t^m = k/m\right)$ is the probability that the remaining inventory at time t is equal to k. For a given N_t^m, let $\xi_1(t), \ldots, \xi_m(t)$ be identically distributed stochastic processes such that $\sum_{i=1}^m \xi_i(t) = \bar{N}_t^m$. By this construction, we know $\mathbb{E}(\xi_i(t) \mid \bar{N}_t^m = k) = \frac{k}{m}$. Therefore, we observe

$$\mathbb{E}\left(\bar{I}_t^m\right) = \sum_{j=0}^{\alpha m} \frac{j}{m} \cdot \mathbb{P}\left(\bar{I}_t^m = \frac{j}{m}\right) = \sum_{j=0}^{\alpha m} \mathbb{E}\left(\xi_i(t) \mid \bar{I}_t^m = \frac{j}{m}\right) \cdot \mathbb{P}\left(\bar{I}_t^m = \frac{j}{m}\right) = \mathbb{E}(\xi_i(t)).$$

We further assume that $\xi_1(t), \xi_2(t), \ldots, \xi_m(t)$ are randomly sampled (without replacement) from a population X, where $X = \{\xi_1(t), \ldots, \xi_M(t)\}$ for some large M. Here, $\xi_1(t), \xi_2(t), \ldots, \xi_M(t)$ are M identically distributed processes such that $\sum_{i=1}^M \xi_i(t) = M\mathbb{E}\left(\bar{I}_t^m\right)$.

Given t, we define the new random variables

$$\eta_i \triangleq \xi_i(t) - \mathbb{E}\left(\bar{I}_t^m\right), \quad \text{for } i = 1, \ldots, m.$$

From our definition of process $\xi_i(t)$ above, we have $\mathbb{E}(\eta_i) = 0$. Let $Y_k \triangleq \sum_{i=1}^k \eta_i$ for $k = 1, \ldots, m$, and let $Y_0 = 0$. Observe that

$$\begin{aligned}
\mathbb{E}\left(\xi_k(t) \mid Y_{k-1}\right) &= \frac{M\mathbb{E}\left(\bar{I}_t^m\right) - \sum_{i=1}^{k-1} \xi_i(t)}{M - (k-1)} \\
&= \frac{M\mathbb{E}\left(\bar{I}_t^m\right) - Y_{k-1} - (k-1)\mathbb{E}(\bar{I}_t^m)}{M - (k-1)} \\
&= \frac{-Y_{k-1}}{M - (k-1)} + \mathbb{E}(\bar{I}_t^m).
\end{aligned} \tag{B.27}$$

Here, the first equality is because $\sum_{i=1}^M \xi_i(t) = M\mathbb{E}(\bar{I}_t^m)$ and $\xi_k(t), \xi_{k+1}(t), \ldots, \xi_M(t)$ are identically distributed. The second equality comes from the definition of $Y_{k-1} = \sum_{i=1}^{k-1} \xi_i(t) - (k-1)\mathbb{E}(\bar{I}_t^m)$.

We further define $Z_k \triangleq \frac{Y_k}{M-k}$, which implies

$$Z_k = \frac{Y_{k-1}}{M-k} + \frac{\eta_k}{M-k} = \frac{M-k+1}{M-k} Z_{k-1} + \frac{\eta_k}{M-k}. \tag{B.28}$$

Now, we analyze the conditional expectation of $\frac{\eta_k}{M-k}$, which is

$$\mathbb{E}\left(\frac{\eta_k}{M-k}|Z_0,\ldots,Z_{k-1}\right)=\mathbb{E}\left(\frac{\eta_k}{M-k}|Y_0,\ldots,Y_{k-1}\right)=\frac{-Y_{k-1}}{(M-(k-1))(M-k)}=-\frac{Z_{k-1}}{M-k}. \tag{B.29}$$

Plugging (B.29) into (B.28), we have

$$\mathbb{E}(Z_k \mid Z_0,\ldots,Z_{k-1})=\frac{M-k+1}{M-k}Z_{k-1}-\frac{Z_{k-1}}{M-k}=Z_{k-1}.$$

Therefore, $\{Z_k, k=0,\ldots,m\}$ is a martingale with respect to the filtration

$$(\sigma(Z_0,\ldots,Z_k))_{k=0,1,2,\ldots,m-1}$$

where $\sigma(Z_0,\ldots,Z_k)$ is the σ-algebra generated by $Z_0,\ldots,Z_k$.

Since $|Z_k-Z_{k-1}|\leq\frac{2\alpha}{M-k+1}$ almost surely, the Hoeffding inequality (applied to martingale) yields

$$\mathbb{P}\left(|Z_m|\geq\frac{m}{M-m}\epsilon\right)\leq 2\exp\left(-\frac{m^2\epsilon^2}{8m\left(1-\frac{m-1}{M}\right)\alpha^2}\right)\quad\text{for any }\epsilon\geq 0 \tag{B.30}$$

where the bound of $\sum_{k=1}^m\frac{4\alpha^2}{(M-k+1)^2}$ comes from Lemma 2.1 in *Serfling* (1974). By integrating (B.30) over $\epsilon\geq 0$, we have

$$\mathbb{E}\left(\frac{M-m}{m}|Z_m|\right)\leq\frac{\sqrt{8\pi\alpha^2}}{\sqrt{m}}.$$

This implies that

$$\mathbb{E}\left|\sum_{j=1}^m\xi_j(t)-m\mathbb{E}\left(\bar{I}_t^m\right)\right|\leq\sqrt{8\pi m\alpha^2}.$$

Because $\bar{I}_t^m=\bar{N}_t^m/m$, the first term on the RHS of (B.26) is $\mathcal{O}(m^{-\frac{1}{2}})$.

For the second term in (B.26), 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,

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

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 III.3, $\mathrm{Var}(D_t^m \mid \mathcal{F}_{t-1}) \le \sigma\lambda^m\left(\bar{N}_{t-1}^m, \alpha m\right) y_t^{\mathsf{OL}}$. Since $\bar{N}_{t-1}^m$ is not random when conditioning on the filtration $\mathcal{F}_{t-1}$, and from (III.4.4) we have

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

where the equality 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

$$\begin{aligned}
&\mathbb{E}\left(\bar{I}_t^m \mid \mathcal{F}_0\right)\\
&\le \frac{1}{2}\mathbb{E}\left[\sqrt{\frac{\sigma\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}}{m}}+\left|\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}-\bar{I}_{t-1}^m\right|-\left(\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}-\bar{I}_{t-1}^m\right)\mid \mathcal{F}_0\right]\\
&\le \frac{1}{2}\left(\Theta(m^{-\frac{1}{2}})+n_{t-1}^{\mathsf{D}}-\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}+n_{t-1}^{\mathsf{D}}-\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}\right)\\
&=\Theta(m^{-\frac{1}{2}})+n_{t-1}^{\mathsf{D}}-\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}. \qquad \text{(B.31)}
\end{aligned}$$

The last inequality comes from the inductive hypotheses (B.24),(B.25). 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

$$\mathbb{E}\left(\mathbb{E}\left(\bar{I}_{t-1}^m-\frac{D_t^m}{m}\mid \mathcal{F}_{t-1}\right)\right)=\mathbb{E}\left(\bar{I}_{t-1}^m-\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}\right)\ge n_{t-1}^{\mathsf{D}}-\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}-\Theta(m^{-\frac{1}{2}}), \qquad \text{(B.32)}$$

where the equality follows from (III.4.4). The inequality is from the inductive hypothesis. Hence, (B.31) and (B.32) 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 (B.26) are both bounded by $\mathcal{O}(m^{-\frac{1}{2}})$, thus giving us (III.4.6) for all t. For a given t, (III.4.7) follows by the Lipschitz continuity of λ and (III.4.6):

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

This concludes the proof. □

B.2.9 Proof of Lemma III.2

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 (III.4.4), (**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 (III.4.8) as

$$\begin{aligned}
&m\left|\mathbb{E}\left[\sum_{t=1}^{T}\left(x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha\right) y_t^{\mathsf{OL}} - x^{-1}(y_t^{\mathsf{D}})\lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}}\right)\right]\right| \\
&\leq m\mathbb{E}\left|\sum_{t=1}^{T}\left(x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha\right) y_t^{\mathsf{OL}} - x^{-1}(y_t^{\mathsf{D}})\lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}}\right)\right| \\
&\leq m\sum_{t=1}^{T}\mathbb{E}\left|x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha\right) y_t^{\mathsf{OL}} - x^{-1}(y_t^{\mathsf{D}})\lambda(n_{t-1}^{\mathsf{D}}, \alpha) y_t^{\mathsf{D}}\right| \\
&= m\sum_{t=1}^{T} x^{-1}(y_t^{\mathsf{D}}) y_t^{\mathsf{D}}\mathbb{E}\left|\lambda\left(\frac{\bar{N}_{t-1}^m}{m}, \alpha\right) - \lambda(n_{t-1}^{\mathsf{D}}, \alpha)\right|.
\end{aligned} \tag{B.33}$$

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 (B.33) is $\mathcal{O}(m^{-\frac{1}{2}})$.

This is true because, from (III.4.7) of Lemma III.1, 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. □

B.2.10 Proof of Theorem III.6

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

$$1-\frac{V^{\mathsf{OL}}(m,T)}{V^{\mathsf{D}}(m,T)} \leq 1-(1-k)\left(1-\frac{C}{2}\sqrt{\frac{\sigma}{m}}-k\right), \tag{B.34}$$

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 (III.2.6), we have

$$V^{\mathsf{OL}}(m,T)=\mathbb{E}\left[\sum_{t=1}^{T}\mathbb{E}\left[x^{-1}\left(y_t^{\mathsf{OL}}\right)\left(D_t^m-[D_t^m-\bar{N}_{t-1}^m]^+\right)\mid\mathcal{F}_{t-1}\right]\right]. \tag{B.35}$$

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 III.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 and from (III.4.4), we get

$$\begin{aligned}
&\mathbb{E}\left[\left[D_t^m-\bar{N}_{t-1}^m\right]^+\mid\mathcal{F}_{t-1}\right]\\
&\leq\frac{\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}+\left(\bar{N}_{t-1}^m-m\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}\right)^2}}{2}-\frac{\left(\bar{N}_{t-1}^m-m\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}\right)}{2}\\
&\leq\frac{1}{2}\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}}+\frac{1}{2}\left|\bar{N}_{t-1}^m-m\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}\right|-\frac{1}{2}\left(\bar{N}_{t-1}^m-m\lambda(\bar{I}_{t-1}^m,\alpha)y_t^{\mathsf{OL}}\right).
\end{aligned}\tag{B.36}$$

Taking the expectation conditioning on $\mathcal{F}_0$ on both sides of (B.36), we have

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[\left[D_t^m-\bar{N}_{t-1}^m\right]^+\mid\mathcal{F}_{t-1}\right]\mid\mathcal{F}_0\right]\\
&\quad\leq\mathbb{E}\left[\frac{1}{2}\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}}\right]+\mathcal{O}(\sqrt{m})\\
&+\frac{1}{2}\left|mn_{t-1}^{\mathsf{D}}-m\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}\right|-\frac{1}{2}\left[mn_{t-1}^{\mathsf{D}}-m\lambda(n_{t-1}^{\mathsf{D}},\alpha)y_t^{\mathsf{D}}\right]\\
&=\mathbb{E}\left[\frac{1}{2}\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^m,\alpha\right)y_t^{\mathsf{OL}}}\right]+\mathcal{O}\left(\sqrt{m}\right).
\end{aligned}\tag{B.37}$$

Here, the first inequality comes from Lemma III.1 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 (III.4.4) and plugging (B.37) into the RHS of (B.35) yields

$$
\begin{aligned}
&V^{\mathsf{OL}}(m,T)\\
&\geq \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\left(m\lambda\left(\bar{I}_{t-1}^{m},\alpha\right) y_t^{\mathsf{OL}} - \frac{1}{2}\sqrt{\sigma m\lambda\left(\bar{I}_{t-1}^{m},\alpha\right) y_t^{\mathsf{OL}}}\right)\right] - \mathcal{O}(\sqrt{m})\\
&= \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right) m\lambda\left(\bar{I}_{t-1}^{m},\alpha\right) y_t^{\mathsf{OL}}\right] - \frac{1}{2}\sqrt{\sigma m}\,\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\sqrt{\lambda\left(\bar{I}_{t-1}^{m},\alpha\right) y_t^{\mathsf{OL}}}\right] - \mathcal{O}(\sqrt{m})\\
&= \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right) m\lambda\left(\bar{I}_{t-1}^{m},\alpha\right) y_t^{\mathsf{OL}}\right]\\
&\quad\times\left(1 - \frac{1}{2}\sqrt{\frac{\sigma}{m}}\underbrace{\frac{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\sqrt{\lambda\left(\bar{I}_{t-1}^{m},\alpha\right) y_t^{\mathsf{OL}}}\right]}{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right)\lambda\left(\bar{I}_{t-1}^{m},\alpha\right) y_t^{\mathsf{OL}}\right]}}_{(**)} - \mathcal{O}\left(1/\sqrt{m}\right)\right).
\end{aligned}
\tag{B.38}
$$

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 (B.38). 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 III.2, we know that the difference between the first term in (B.38) 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,

$$
\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(y_t^{\mathsf{OL}}\right) m\lambda\left(\bar{I}_{t-1}^{m},\alpha\right) y_t^{\mathsf{OL}}\right) \geq V^{\mathsf{D}}(m,T)(1-k),
\tag{B.39}
$$

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 (B.38). Note that from Cauchy-Swartz inequality, the numerator of $(**)$ is bounded above by

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

where the first inequality comes from Assumption III.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 (B.39).

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

$$(**) \leq \sqrt{\frac{Tx^{-1}(0)}{V^{\mathsf{D}}(T)}}\Theta(1) \triangleq C. \tag{B.40}$$

Finally, we take (B.39) and (B.40) into (B.38), 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. □

B.2.11 Proof of Lemma III.3

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

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

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

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

where $\bar{y}$ is defined in Lemma B.1(ii). When $n \in S_1$, $\mathbf{y}^{\mathsf{CE}}(n,t)$ has bounded derivative w.r.t. n because of Lemma B.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 $\mathbf{y}^{\mathsf{CE}}(n,t)$ w.r.t. n is bounded for $n \in [0,\alpha]$. By definition, $\mathbf{y}^{\mathsf{CE}}(n,t) = y_0^{\mathsf{D}}(n, T-t+1)$ 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 (B.7).

By Claim B.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

$$\frac{\partial}{\partial y}\left(x^{-1}(y)y\right)\Big|_{y=\bar{y}_{t,n}} = \frac{\partial V^{\mathsf{D}}(T-t; n', \alpha)}{\partial n'}\Big|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}}. \tag{B.41}$$

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

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

From Lemma B.1(iii), the derivative of $\mathbf{y}^{\mathsf{CE}}(n,t)$ w.r.t. n is bounded when $n \in S_1'$. When $n \in S_2'$, the derivative of $\mathbf{y}^{\mathsf{CE}}(n,t) = \bar{y}_{t,n}$ w.r.t. n. To this, differentiate (B.41) 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

$$\frac{\partial \bar{y}_{t,n}}{\partial n}\left(x^{-1}(y)y\right)''\Big|_{y=\bar{y}_{t,n}} = \left(1 - \lambda_1(n,\alpha)\bar{y}_{t,n} - \lambda(n,\alpha)\frac{\partial \bar{y}_{t,n}}{\partial n}\right)\frac{\partial^2 V^{\mathsf{D}}(T-t; n', \alpha)}{\partial n'^2}\Big|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}}. \tag{B.42}$$

Rearranging terms in (B.42) yields the following relationship:

$$\left|\frac{\partial \bar{y}_{t,n}}{\partial n}\right| = \left| \frac{(1-\lambda_1(n,\alpha)\bar{y}_{t,n}) \left.\dfrac{\partial^2 V^{\mathsf{D}}(T-t;n',\alpha)}{\partial n'^2}\right|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}}}{(x^{-1}(y)y)''\Big|_{y=\bar{y}_{t,n}} + \lambda(n,\alpha)\left.\dfrac{\partial^2 V^{\mathsf{D}}(T-t;n',\alpha)}{\partial n'^2}\right|_{n'=n-\lambda(n,\alpha)\bar{y}_{t,n}}} \right|. \tag{B.43}$$

The term on the RHS of (B.43) is bounded (i.e., the denominator is nonzero) because $r''(y) < 0$ is defined for $y \in [0,1]$ according to Lemma B.1(i), $\partial^2 V^{\mathsf{D}}(T-t;n',\alpha)/\partial n'^2 < 0$ is defined for $n' \in [0,1]$ (Theorem III.2(ii)), and $\lambda(n,\alpha)$ is continuous differentiable for $n \in [0,\alpha]$ and finite $\alpha \geq 0$. This concludes our proof. □

B.2.12 Lemma B.4 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 $\hat{N}^m = (\hat{N}^m_0, \hat{N}^m_1 \ldots, \hat{N}^m_T)$, where $\hat{N}^m_0 = \alpha m$. 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 $u = \hat{N}^m_{t-1}$. (As we discussed in Section III.4.2, ($\mathbf{D}_m$) is the scaled version of ($\mathbf{D}$). Hence, by the inventory constraint ($\mathbf{D}$b), the period t conditional expected demand under policy CE-CL would never exceed N^m_{t-1}.)

Lemma B.4 (Convergence of remaining inventory and SIS)**.** If $n^{\mathsf{D}} = (n^{\mathsf{D}}_1, \ldots, n^{\mathsf{D}}_T)$ is the solution to ($\mathbf{D}$) when $u = \alpha$, then the following hold:

$$\mathbb{E}\left|\frac{\hat{N}^m_t}{m} - n^{\mathsf{D}}_t\right| = \mathcal{O}\left(1/\sqrt{m}\right), \quad \text{for all } t = 1, \ldots, T \tag{B.44}$$

$$\mathbb{E}\left|\lambda\left(\frac{\hat{N}^m_t}{m}, \alpha\right) - \lambda\left(n^{\mathsf{D}}_t, \alpha\right)\right| = \mathcal{O}\left(1/\sqrt{m}\right), \quad \text{for all } t = 1, \ldots, T \tag{B.45}$$

Proof. The proof is analogous to that of Lemma III.1 in Appendix B.2.8. We start by defining the sequence of random variables $(\hat{I}^m_0, \hat{I}^m_1, \ldots, \hat{I}^m_T)$, where $\hat{I}^m_t = \hat{N}^m_t/m$ is the normalized remaining inventory at time t under the closed-loop policy $\mathbf{y}^{\mathsf{CE}}$ when the initial inventory and the expected demand are scaled by m. Note that $\hat{I}^m_0 = \alpha$.

We will prove the lemma by induction. The base case is $t = 0$, where we note that $\hat{I}^m_0 = n^{\mathsf{D}}_0 = \alpha$, and hence $\lambda(\hat{I}^m_0, \alpha) = \lambda(n^{\mathsf{D}}_0, \alpha) = \lambda(\alpha, \alpha)$. Therefore, (B.44) and (B.45) hold for $t = 0$. For the induction step, we assume that (B.44) and (B.45) hold for $t-1$.

Specifically,

$$\mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| = \mathcal{O}\left(1/\sqrt{m}\right), \tag{B.46}$$

$$\mathbb{E}\left|\lambda\left(\hat{I}^m_{t-1}, \alpha\right) - \lambda\left(n^{\mathsf{D}}_{t-1}, \alpha\right)\right| = \mathcal{O}\left(1/\sqrt{m}\right), \tag{B.47}$$

We will show these properties (B.44),(B.45) hold for t.

To prove (B.44) for t, notice that by adding and subtracting $\mathbb{E}(\hat{I}^m_t)$,

$$\mathbb{E}\left|\hat{I}^m_t - n^{\mathsf{D}}_t\right| = \mathbb{E}\left|\hat{I}^m_t - \mathbb{E}\left(\hat{I}^m_t\right) + \mathbb{E}\left(\hat{I}^m_t\right) - n^{\mathsf{D}}_t\right| \le \mathbb{E}\left|\hat{I}^m_t - \mathbb{E}\left(\hat{I}^m_t\right)\right| + \left|\mathbb{E}\left(\hat{I}^m_t\right) - n^{\mathsf{D}}_t\right|. \tag{B.48}$$

We will show that both terms in the right side of (B.48) are $\mathcal{O}(1/\sqrt{m})$.

Following the similar argument from the proof of Lemma III.1 in Appendix B.2.8 until (B.30), we have the first term on the RHS of (B.48) is $\mathcal{O}(1/\sqrt{m})$. For the second term in (B.48), we want to bound the difference between $\mathbb{E}(\hat{I}^m_t)$ and n^{D}_t. From the definition of $\hat{I}^m_t$, we know

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

Note $\mathbb{E}(D^m_t \mid \mathcal{F}_{t-1}) = \lambda^m\left(\hat{N}^m_{t-1}, \alpha m\right)\mathbf{y}^{\mathsf{CE}}\left(\hat{N}^m_{t-1}, t\right)$ and, by Assumption III.3, we also have a bound on the variance $\mathrm{Var}(D^m_t \mid \mathcal{F}_{t-1}) \le \sigma\lambda^m(\hat{N}^m_{t-1}, \alpha m)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t)$. Therefore since $\hat{N}^m_{t-1}$ is not random when conditioning on the filtration $\mathcal{F}_{t-1}$, and using the Scarf bound and (III.4.4), we have

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

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 $\mathbf{y}^{\mathsf{CE}}$ always satisfies $\hat{N}^m_{t-1} - \lambda(\hat{N}^m_{t-1}, \alpha m)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t) \ge 0$ 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

$$\mathbb{E}(\hat{I}^m_t \mid \mathcal{F}_{t-1}) \le \hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t) + \Theta(m^{-\frac{1}{2}}). \tag{B.49}$$

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

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

We also have a lower bound from the following arguments:

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

where the last relationship uses (III.4.4). Hence,

$$0 \le \mathbb{E}\left(\hat{I}^m_t - \hat{I}^m_{t-1} + \lambda(\hat{I}^m_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t)\right) \le \Theta(1/\sqrt{m}). \tag{B.50}$$

This implies that

$$\begin{aligned}
\left|\mathbb{E}(\hat{I}^m_t) - n^{\mathsf{D}}_t\right| &= \left|\mathbb{E}(\hat{I}^m_t) - n^{\mathsf{D}}_{t-1} + \lambda(n^{\mathsf{D}}_{t-1}, \alpha)y^{\mathsf{D}}_t\right| \\
&\le \left|\mathbb{E}\left(\hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t)\right) - n^{\mathsf{D}}_{t-1} + \lambda(n^{\mathsf{D}}_{t-1}, \alpha)y^{\mathsf{D}}_t\right| + \Theta(1/\sqrt{m}) \qquad \text{(B.51)} \\
&\le \mathbb{E}\left|\hat{I}^m_{t-1} - \lambda(\hat{I}^m_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t) - n^{\mathsf{D}}_{t-1} + \lambda(n^{\mathsf{D}}_{t-1}, \alpha)y^{\mathsf{D}}_t\right| + \Theta(1/\sqrt{m}) \qquad \text{(B.52)} \\
&\le \mathbb{E}|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}| + \mathbb{E}|\lambda(\hat{I}^m_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t) - \lambda(n^{\mathsf{D}}_{t-1}, \alpha)y^{\mathsf{D}}_t| + \Theta(1/\sqrt{m}) \qquad \text{(B.53)} \\
&\le \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| + \underbrace{\mathbb{E}\left|\lambda(\hat{I}^m_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t) - \lambda(n^{\mathsf{D}}_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t)\right|}_{(*)} \\
&\quad + \underbrace{\mathbb{E}\left|\lambda(n^{\mathsf{D}}_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t) - \lambda(n^{\mathsf{D}}_{t-1}, \alpha)y^{\mathsf{D}}_t\right|}_{(**)} + \Theta(1/\sqrt{m}), \qquad \text{(B.54)}
\end{aligned}$$

where (B.51) follows from (B.50), (B.52) is from Jensen's inequality, (B.53) is from triangle inequality and monotonicity of expectation, (B.54) is derived by subtracting and adding $\lambda(n^{\mathsf{D}}_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t)$ 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 III.2(vi)), so there exists a C_λ such that $|\lambda(n, \alpha) - \lambda(n', \alpha)| \le C_\lambda |n - n'|$ for all n, n', and fixed α. Also, we know

$\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t) \leq 1$ by Assumption III.2(i). Therefore,

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

To analyze the bound for $(**)$, we know from Lemma III.3 that $\mathbf{y}^{\mathsf{CE}}(n, t)$ is Lipschitz continuous in n with some Lipschitz constant C_y. Furthermore, observe that $\mathbf{y}^{\mathsf{CE}}(mn^{\mathsf{D}}_t, t) = y^{\mathsf{D}}_t$. Another important property of $\mathbf{y}^{\mathsf{CE}}$ we need is that $\mathbf{y}^{\mathsf{CE}}(mn, t; m\alpha)$ under initial inventory is $m\alpha$ is the same as $\mathbf{y}^{\mathsf{CE}}(n, t; \alpha)$ under initial inventory is α. This is because $\mathbf{y}^{\mathsf{CE}}$ 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 (III.4.4). Therefore,

$$\begin{aligned}
(**) &= \mathbb{E}\left|\lambda(n^{\mathsf{D}}_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1}, t; m\alpha) - \lambda(n^{\mathsf{D}}_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(mn^{\mathsf{D}}_t, t; m\alpha)\right| \\
&= \mathbb{E}\left|\lambda(n^{\mathsf{D}}_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{I}^m_{t-1}, t; \alpha) - \lambda(n^{\mathsf{D}}_{t-1}, \alpha)\mathbf{y}^{\mathsf{CE}}(n^{\mathsf{D}}_t, t; \alpha)\right| \\
&\leq \bar{\lambda} C_y \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| \qquad \text{(B.55)}
\end{aligned}$$

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

$$\begin{aligned}
\left|\mathbb{E}(\hat{I}^m_t) - n^{\mathsf{D}}_t\right| &\leq \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| + 1 \cdot C_\lambda \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| + \bar{\lambda} C_y \mathbb{E}\left|\hat{I}^m_{t-1} - n^{\mathsf{D}}_{t-1}\right| + \Theta(1/\sqrt{m}), \\
&= \mathcal{O}\left(1/\sqrt{m}\right), \qquad \text{(B.56)}
\end{aligned}$$

where (B.56) comes from the inductive hypothesis (B.46).

Therefore, we can conclude that the RHS two terms of (B.48) are both bounded by $\mathcal{O}\left(1/\sqrt{m}\right)$, thus giving us (B.44) for all t by induction. For a given t, (B.45) follows by the Lipschitz continuity of λ and (B.44):

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

This concludes the proof. □

B.2.13 Lemma B.5 and proof

An important implication of Lemma B.4 is that the intensity policy $\mathbf{y}^{\mathsf{CE}}$ converges to the deterministic sequence $\mathbf{y}^{\mathsf{D}}$ since, with Lemma III.3, we know that $\mathbf{y}^{\mathsf{CE}}$ is Lipschitz continuous. These properties allow us show that the *uncensored* expected revenue under $\mathbf{y}^{\mathsf{CE}}$ has a gap from $V^{\mathsf{D}}(m, T)$ that is $\mathcal{O}(\sqrt{m})$. This is formalized in the lemma below.

Lemma B.5 (Convergence of uncensored revenue)**.**

$$\left|\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda^m(\hat{N}^m_{t-1},\alpha m)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right) - V^{\mathsf{D}}(m,T)\right| = \mathcal{O}\left(\sqrt{m}\right). \tag{B.57}$$

Proof. By definition of V^{D} and from property (III.4.4) 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 (B.57) as

$$\begin{aligned}
&m\left|\mathbb{E}\left[\sum_{t=1}^{T}\left(x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t) - x^{-1}(y^{\mathsf{D}}_t)\lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right)\right]\right| \\
&\le m\mathbb{E}\left|\sum_{t=1}^{T}\left(x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t) - x^{-1}(y^{\mathsf{D}}_t)\lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right)\right| \\
&\le m\sum_{t=1}^{T}\mathbb{E}\left|x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t) - x^{-1}(y^{\mathsf{D}}_t)\lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right|. \qquad \text{(B.58)}
\end{aligned}$$

Here, the first inequality comes from $|\mathbb{E}X| \le \mathbb{E}|X|$ 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 (B.58) is $\mathcal{O}\left(1/\sqrt{m}\right)$.

Note that for any t,

$$\begin{aligned}
&\mathbb{E}\left|x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t) - x^{-1}(y^{\mathsf{D}}_t)\lambda(n^{\mathsf{D}}_{t-1},\alpha)y^{\mathsf{D}}_t\right| \\
&= \mathbb{E}\left|r\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha) - r(y^{\mathsf{D}}_t)\lambda(n^{\mathsf{D}}_{t-1},\alpha)\right|, \qquad \text{(B.59)}
\end{aligned}$$

where $r(y) = x^{-1}(y)y$ is the per-period revenue rate. Our goal is to show that (B.59) is $\mathcal{O}(1/\sqrt{m})$.

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

$$|r(y) - r(y')| \le C_r|y - y'|. \tag{B.60}$$

Additionally, $r(y) \le \bar{f} = r(\bar{y})$ where $\bar{y}$ is defined in Lemma B.1(ii). Hence, if we subtract and add the term $r(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t))\lambda(n^{\mathsf{D}}_{t-1},\alpha)$ inside the absolute value in (B.59), by triangle inequality, (B.59) is upper bounded by

$$\begin{aligned}
&\mathbb{E}\left| r\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha) - r\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(n^{\mathsf{D}}_{t-1},\alpha)\right| \\
&\quad + \mathbb{E}\left| r\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(n^{\mathsf{D}}_{t-1},\alpha) - r(y^{\mathsf{D}}_t)\lambda(n^{\mathsf{D}}_{t-1},\alpha)\right| \\
&\leq \bar{f}\ \mathbb{E}\left|\lambda(\hat{I}^m_{t-1},\alpha) - \lambda(n^{\mathsf{D}}_{t-1},\alpha)\right| + \bar{\lambda}C_r\ \mathbb{E}\left|\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t) - y^{\mathsf{D}}_t\right|,
\end{aligned} \tag{B.61}$$

where the second term of (B.61) comes from (B.60) and Assumption III.2(v). Hence, it suffices to show the two terms in (B.61) are bounded by $\mathcal{O}(1/\sqrt{m})$. This is true because, from (B.45) of Lemma B.4, for any t,

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

Moreover, by definition, $\mathbf{y}^{\mathsf{CE}}$ results from re-optimizing the deterministic equivalent at each time period, hence we have that $\mathbf{y}^{\mathsf{CE}}(mn^{\mathsf{D}}_{t-1},t) = y^{\mathsf{D}}_t$. We use the property of $\mathbf{y}^{\mathsf{CE}}$ that $\mathbf{y}^{\mathsf{CE}}(mn,t;m\alpha)$ under initial inventory is $m\alpha$ is the same as $\mathbf{y}^{\mathsf{CE}}(n,t;\alpha)$ under initial inventory is α. This is because $\mathbf{y}^{\mathsf{CE}}$ solves optimization model (**D**) where the optimal deterministic intensity solution is invariant under scaling since (III.4.4) implies that, for any $n \in [0,\alpha]$, $\lambda^m(mn,m\alpha) = m\lambda(n,\alpha)$. Therefore,

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

where the inequality is from Lemma III.3, and the last equality is from (B.44) of Lemma B.4. This concludes the proof. □

B.2.14 Proof of Theorem III.7

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

$$1 - \frac{V^{\mathsf{CE}}(m,T)}{V^{\mathsf{D}}(m,T)} \leq 1 - (1-k)\left(1 - \frac{C}{2}\sqrt{\frac{\sigma}{m}}\right), \tag{B.62}$$

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

Recall $\hat{N}^m = (\hat{N}^m_0, \ldots, \hat{N}^m_T)$ is the stochastic sequence of remaining inventories under

$\mathbf{y}^{\mathsf{CE}}$, where initial inventory is $\hat{N}_0^m = \alpha m$. Then from (III.2.6), we have

$$V^{\mathsf{CE}}(m,T) = \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{E}\left[x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t)\right)\left(D_t^m - [D_t^m - \hat{N}_{t-1}^m]^+\right) \mid \mathcal{F}_{t-1}\right]\right]. \tag{B.63}$$

We next define 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, $\mathbb{E}(D_t^m \mid \mathcal{F}_{t-1}) = m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t)$ and, by Assumption III.3,

$$\mathrm{Var}(D_t^m \mid \mathcal{F}_{t-1}) \le \sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t).$$

Therefore, by the Scarf bound and from (III.4.4) we have

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

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

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

then we have the following:

$$\begin{aligned}
&\mathbb{E}\left[[D_t^m - \hat{N}_{t-1}^m]^+ \mid \mathcal{F}_{t-1}\right] \\
&\le \frac{1}{2}\frac{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t)}{\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t) + (\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t))^2} + \left(\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t)\right)} \\
&\le \frac{1}{2}\frac{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t)}{\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t) + (\hat{N}_{t-1}^m - m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t))^2}} \\
&\le \frac{1}{2}\sqrt{\sigma m\lambda(\hat{I}_{t-1}^m, \alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t)}.
\end{aligned} \tag{B.64}$$

The second inequality is because, conditional on $\mathcal{F}_{t-1}$, $\hat{N}_{t-1}^m - \lambda(\hat{N}_{t-1}^m, \alpha m)\mathbf{y}^{\mathsf{CE}}(\hat{N}_{t-1}^m, t) \ge 0$. 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 (B.64) into the RHS of (B.63), we observe that

$$\begin{aligned}
V^{\mathsf{CE}}(m,T) &\geq \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\left(m\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t) - \frac{1}{2}\sqrt{\sigma m\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)}\right)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right) m\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right] \\
&\qquad - \frac{1}{2}\sqrt{\sigma m}\,\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\sqrt{\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)}\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right) m\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right] \\
&\quad \times \left(1 - \frac{1}{2}\sqrt{\frac{\sigma}{m}}\underbrace{\frac{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\sqrt{\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)}\right]}{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right]}}_{(**)}\right). \qquad \text{(B.65)}
\end{aligned}$$

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 (B.65). Note that $\mathbf{y}^{\mathsf{CE}}$ does not scale with m since it it is constructed from the intensity solution of (**D**) which is scale-invariant due to property (III.4.4) of λ. From Lemma B.5, we know that the difference between the first term in (B.65) 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,

$$\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right) m\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right] \geq V^{\mathsf{D}}(m,T)(1-k) \qquad \text{(B.66)}$$

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 (B.65). From Cauchy-Swartz inequality, the numerator of $(**)$ is bounded above by

$$\begin{aligned}
&\mathbb{E}\left[\sqrt{\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)}\sqrt{\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)}\right] \\
&\leq \mathbb{E}\left[\sqrt{\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)}\right]\sqrt{Tx^{-1}(0)} \\
&\leq \sqrt{\mathbb{E}\left[\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right)\lambda(\hat{I}^m_{t-1},\alpha)\mathbf{y}^{\mathsf{CE}}(\hat{N}^m_{t-1},t)\right]}\sqrt{Tx^{-1}(0)},
\end{aligned}$$

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

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

where the last inequality comes from (B.66).

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

$$(**) \leq \sqrt{\frac{Tx^{-1}(0)}{V^{\mathsf{D}}(T)}}\Theta(1) \triangleq C. \tag{B.67}$$

Finally, we take (B.66) and (B.67) into (B.65), resulting in

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

This completes the proof. □

B.3 Section III.5 proofs

B.3.1 Proof of Theorem III.8

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 III.1 to III.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 (B.65) and (B.67), we know

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

Note that Lemmas III.2 and B.5 implies that

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

where $k=\mathcal{O}(1/\sqrt{m})$ and $k \geq 0$. Therefore, subtracting both sides of (B.1) by $c\alpha^{\mathsf{CE}}m$, and using (B.2), we have

$$\underbrace{V^{\alpha^{\mathsf{CE}},\mathbf{y}^{\mathsf{CE}}}(m,T)-c\alpha^{\mathsf{CE}}m}_{Q^{\mathsf{CE}}(m,T)} \geq m\left(V^{\mathsf{D},\alpha^{\mathsf{CE}}}(T)-k\right)\left(1-\frac{1}{2}\sqrt{\frac{\sigma}{m}}C\right)-c\alpha^{\mathsf{CE}}m. \tag{B.3}$$

Now we analyze the RHS of (B.3) to connect it to $Q^{*}(m,T)$. Define $k_1=\frac{1}{2}\sqrt{\frac{\sigma}{m}}C$ where C is defined in (B.67) with $\alpha=\alpha^{\mathsf{CE}}$.

Factoring out $m(1-k_1)$ in the RHS of (B.3) results in

$$\begin{aligned}
& m(1-k_1)\left(V^{\mathsf{D},\alpha^{\mathsf{CE}}}(T)-k-\frac{c\alpha^{\mathsf{CE}}}{1-k_1}\right) \\
&= m(1-k_1)\left(\underbrace{V^{\mathsf{D},\alpha^{\mathsf{CE}}}(T)-c\alpha^{\mathsf{CE}}}_{Q^{\mathsf{D},\alpha^{\mathsf{CE}}}(T)}-k+c\alpha^{\mathsf{CE}}-\frac{c\alpha^{\mathsf{CE}}}{1-k_1}\right) && \text{subtracting and adding } c\alpha^{\mathsf{CE}} \\
&\geq m(1-k_1)\left(\underbrace{V^{\mathsf{D},\alpha^{*}}(T)-c\alpha^{*}}_{Q^{\mathsf{D},\alpha^{*}}(T)}-k-c\alpha^{\mathsf{CE}}\frac{k_1}{1-k_1}\right) && \text{definition of } \alpha^{\mathsf{CE}} \text{ so } Q^{\mathsf{D},\alpha^{\mathsf{CE}}}(T)\geq Q^{\mathsf{D},\alpha^{*}}(T) \\
&= (1-k_1)\left(V^{\mathsf{D},\alpha^{*}}(m,T)-c\alpha^{*}m-mk-c\alpha^{\mathsf{CE}}m\frac{k_1}{1-k_1}\right) && \text{multiplying } m \text{ inside} \\
&\geq (1-k_1)\left(V^{\alpha^{*},\mathbf{y}^{*}}(m,T)-c\alpha^{*}m-mk-c\alpha^{\mathsf{CE}}m\frac{k_1}{1-k_1}\right) && \text{from Proposition III.1} \\
&= (1-k_1)\left(\underbrace{V^{\alpha^{*},\mathbf{y}^{*}}(m,T)-c\alpha^{*}m}_{Q^{*}(m,T)}-(m+c\alpha^{\mathsf{CE}}m)k_2\right)
\end{aligned} \tag{B.4}$$

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 (B.3) and the RHS of (B.4) 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 (III.5.1), 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. □

B.3.2 Proof of Proposition III.2

Since it is not possible to characterize the exact revenue difference between the optimal and the fixed price policy, to prove Proposition III.2, we utilize the bound established by V^{D}. To see this, an implication of our results in Section III.3 is $0 \leq V^{\mathsf{D}}(m,T) - V^*(m,T) \leq \mathcal{O}(\sqrt{m})$ (Proposition III.1, Theorem III.7). 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

$$V^{\mathsf{D},\alpha}(m,T) - V^{\mathsf{SP},\alpha}(m,T) = \Omega(m), \tag{B.5}$$

then this establishes the first statement in Proposition III.2. Note that this also proves the second statement since the profit loss of the fixed price policy $(\alpha^{\mathsf{SP}}, \mathbf{y}^{\mathsf{SP}})$ is bounded below by the revenue loss of $\mathbf{y}^{\mathsf{SP}}$ with $\alpha = \alpha^{\mathsf{SP}}$.

We need two key results to prove (B.5). The first key result in establishing (B.5) 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 (III.5.3) when the initial inventory is αm. This is formalized in the following lemma (whose proof is in Appendix B.3.3) that states that the difference grows at a linear rate in m.

Lemma B.6 (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 $V^{\mathsf{D},\alpha}(m,T) - V^{\mathsf{D}',\alpha}(m,T) = \Theta(m)$.

Condition (i) of Lemma B.6 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 III.3, is that the gap between the expected revenue $V^{\mathsf{SP},\alpha}(m,T)$ and the deterministic revenue $V^{\mathsf{D}',\alpha}(m,T)$ is $\mathcal{O}(\sqrt{m})$. Note that by $V^{\mathsf{SP},\alpha}(m,T)$ we mean the expected revenue of the fixed price policy under the stochastic problem. The proof is in Appendix B.3.6.

Lemma B.7. For a fixed $\alpha \geq 0$,

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

Now we are ready to prove the proposition.

Proof of Proposition III.2. From the definition that $Q^*(m,T)$ is the optimal profit, we know $Q^*(m,T) \geq V^{*,\alpha^{\mathsf{SP}}}(m,T) - m\alpha^{\mathsf{SP}}c$. Then,

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

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

We know that $V^{*,\alpha}(m,T)$ is bounded below by $V^{\mathsf{CE},\alpha}(m,T)$. Hence, by Theorem III.7, we have that $V^{*,\alpha}(m,T) \geq V^{\mathsf{D},\alpha}(m,T) - \mathcal{O}(\sqrt{m})$. This and Lemma B.7 result in

$$V^{*,\alpha}(m,T) - V^{\mathsf{SP},\alpha}(m,T) \geq V^{\mathsf{D},\alpha}(m,T) - \mathcal{O}(\sqrt{m}) - V^{\mathsf{D}',\alpha}(m,T) - \mathcal{O}(\sqrt{m}). \quad \text{(B.6)}$$

Moreover, according to Lemma B.6, we know the RHS of (B.6) equals to $\Theta(m) - \mathcal{O}(\sqrt{m})$, which is $\Omega(m)$. This concludes the proof. $\square$

B.3.3 Proof of Lemma B.6

Proof. Consider an arbitrary $\alpha \geq 0$ satisfying the conditions of the lemma. Recall the definition $R^{\mathsf{D}}(u,T)$ in (B.6), where $V^{\mathsf{D}}(T) = R^{\mathsf{D}}(\alpha, T)$.

Due to condition (ii) of the lemma and from (III.5.2), we have that $y^{\mathsf{SP}} = \bar{y}$. 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 (B.6). Note that $R^{\mathsf{D},y}(u,T)$ is the objective in (B.6). 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

$$V^{\mathsf{D}}(T) - V^{\mathsf{D}'}(T) = \underbrace{R^{\mathsf{D},y_1^{\mathsf{D}}}(\alpha,T) - R^{\mathsf{D},\bar{y}}(\alpha,T)}_{(a)} + \underbrace{R^{\mathsf{D},\bar{y}}(\alpha,T) - R^{\mathsf{D}',\bar{y}}(\alpha,T)}_{(b)}. \tag{B.7}$$

In (B.7), $(b) \geq 0$ because

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

since $R^{\mathsf{D}}(\cdot,\cdot) = V^{\mathsf{D}}(\cdot,\cdot)$ defined in (**D**), 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**)). Therefore, the RHS of (B.7) is lower bounded by (a).

Because $R^{\mathsf{D},y}$ is strictly concave in y (Claim B.2) and since $y_1^{\mathsf{D}} > 0$ (Theorem III.4), then we know

$$\frac{\partial R^{\mathsf{D},y}(\alpha,T)}{\partial y}\Big|_{y\,=\,y_1^{\mathsf{D}}} = \underbrace{\frac{\partial}{\partial y} x^{-1}(y)\lambda(\alpha,\alpha)y\Big|_{y\,=\,y_1^{\mathsf{D}}}}_{(c)} + \underbrace{\frac{\partial}{\partial y} R^{\mathsf{D}}(\alpha - \lambda(\alpha,\alpha)y, T-1)\Big|_{y\,=\,y_1^{\mathsf{D}}}}_{(d)} = 0. \tag{B.8}$$

Condition (i) of Lemma B.6 states that $(d) \neq 0$ which, combined with (B.8), implies that $(c) \neq 0$. Since $\bar{y}$ is the unique value that can make $\frac{\partial}{\partial y} x^{-1}(y)\lambda(\alpha,\alpha)y$ equal to zero (Lemma B.1(ii)), we conclude $y_1^{\mathsf{D}} \neq \bar{y}$. Therefore, by the mean value theorem, there exists a $y' \in \left(\min\{\bar{y}, y_1^{\mathsf{D}}\}, \max\{\bar{y}, y_1^{\mathsf{D}}\}\right)$ such that

$$(a) = R^{\mathsf{D},y_1^{\mathsf{D}}}(\alpha,T) - R^{\mathsf{D},\bar{y}}(\alpha,T) = \frac{\partial R^{\mathsf{D},y}(\alpha,T)}{\partial y}\Big|_{y=y'} (y_1^{\mathsf{D}} - \bar{y}). \tag{B.9}$$

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 (B.9) is nonzero because $y' \neq y_1^{\mathsf{D}}$ and y_1^{D} is the unique maximizer of $R^{\mathsf{D},y}(\alpha,T)$ (Lemma B.1(ii)). Further, since $y_1^{\mathsf{D}} \neq \bar{y}$, we have that $(a) > 0$. Hence, $V^{\mathsf{D}}(T) - V^{\mathsf{SP}}(T) > 0$. This implies that $V^{\mathsf{D}}(m,T) - V^{\mathsf{SP}}(m,T) = m\left(V^{\mathsf{D}}(T) - V^{\mathsf{SP}}(T)\right) = \Theta(m)$. This concludes our proof. □

B.3.4 Corollary B.1 and proof

Corollary B.1. Given $\alpha \geq 0$, let $(N_0^m, N_1^m, \ldots, N_T^m)$ denote the sequence of stochastic remaining inventory under policy $\mathbf{y}^{\mathsf{SP}}$ with $N_0^m = \alpha m$. Define $I_t^m \triangleq N_t^m/m$. Let $n^{\mathsf{D}'} = (n_0^{\mathsf{D}'}, \ldots, n_T^{\mathsf{D}'})$ be the deterministic sequence of remaining inventory when fixing $y = (y^{\mathsf{SP}}, \ldots, y^{\mathsf{SP}})$ with initial inventory α. Then the following hold:

$$\mathbb{E}\left|I_t^m - n_t^{\mathsf{D}'}\right| = \mathcal{O}\left(1/\sqrt{m}\right)$$

and

$$\mathbb{E}\left|\lambda\left(I_t^m,\alpha\right) - \lambda\left(n_t^{\mathsf{D}'},\alpha\right)\right| = \mathcal{O}\left(1/\sqrt{m}\right).$$

Proof. The only difference between Corollary B.1 and Lemma B.4 is the gap between the stochastic intensity sequence and the deterministic intensity sequence. In Lemma B.4 (using the notation in the proof of Lemma B.4), we apply $\mathbf{y}^{\mathsf{CE}}$ to the stochastic problem and accordingly get normalized inventory $(\hat{I}_t^m)_t$; and we apply y^{D} to the deterministic problem and accordingly have n^{D}. However, in Corollary B.1, we apply $(y^{\mathsf{SP}}, \ldots, y^{\mathsf{SP}})$ to the stochastic problem and accordingly get normalized inventory $(I_t^m)_t$; and we apply the same $(y^{\mathsf{SP}}, \ldots, y^{\mathsf{SP}})$ to the deterministic problem and accordingly have $n^{\mathsf{D}'}$. As a result, the key difference between the proofs of Lemma B.4 and Corollary B.1 is the logic to have the same $(**)$ in (B.54) upper bounded by (B.55). Note that the definition of $\mathbf{y}^{\mathsf{SP}}$ in (III.5.2) also guarantees that inventory constraint is satisfied in expectation, so the logic in the proof stays the same as Lemma B.4.

In Lemma B.4, (using the notation in the proof of Lemma B.4) we have the gap between $\mathbf{y}^{\mathsf{CE}}$ and y^{D} is

$$\mathbb{E}\left|\mathbf{y}^{\mathsf{CE}}\left(\hat{N}_{t-1}^{m}, t\right) - y_t^{\mathsf{D}}\right| \leq \bar{\lambda} C_y \mathbb{E}\left|\hat{I}_{t-1}^{m} - n_{t-1}^{\mathsf{D}}\right| = \mathcal{O}\left(1/\sqrt{m}\right). \tag{B.10}$$

Note that (B.10) is the key to have $(**) \leq$ (B.55) in the proof of Lemma B.4. To get (B.10), the crucial part is the Lipschitz continuity of policy y^{CE} proved in Lemma III.3. Therefore, in Corollary B.1, if we also have the gap between y sequences applied to the stochastic and deterministic problems is $\mathcal{O}(1/\sqrt{m})$, then we are done. In fact, for Corollary B.1, we apply the same sequence $(y^{\mathsf{SP}}, \ldots, y^{\mathsf{SP}})$ to both stochastic and deterministic problems, so clearly

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

thus is $\mathcal{O}(1/\sqrt{m})$. Therefore, we get the same bound as (B.55) in the proof of Lemma B.4. Then, Corollary B.1 holds by applying the same logic as the proof of Lemma B.4.

□

B.3.5 Corollary B.2 and proof

Corollary B.2. Given $\alpha \geq 0$, let $(N_0^m, N_1^m, \ldots, N_T^m)$ denote the sequence of stochastic remaining inventory under policy $\mathbf{y}^{\mathsf{SP}}$ with $N_0^m = \alpha m$. Define $I_t^m \triangleq N_t^m/m$. Then,

$$\left|\mathbb{E}\left(\sum_{t=1}^{T} x^{-1}\left(\mathbf{y}^{\mathsf{SP}}(N_{t-1}^m, t)\right) \lambda\left(I_{t-1}^m, \alpha\right) \mathbf{y}^{\mathsf{SP}}(N_{t-1}^m, t)\right) - V^{\mathsf{SP}}(\alpha, T)\right| = \mathcal{O}\left(1/\sqrt{m}\right).$$

Proof. Similar to the proof of Corollary B.1 (see Appendix B.3.4), the only difference between Corollary B.2 and Lemma B.5 is the gap between the stochastic intensity sequence and the deterministic intensity sequence. In Lemma B.5 (using the notation in the proof of Lemma B.5), we apply $\mathbf{y}^{\mathsf{CE}}$ to the stochastic problem and accordingly get the remaining inventory $(\hat{N}_t^m)_{t=0}^T$ and the expected revenue

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

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 B.1, we apply $(y^{\mathsf{SP}}, \ldots, y^{\mathsf{SP}})$ to the stochastic problem and accordingly get the normalized inventory $(I_t^m)_t$ and the expected revenue

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

and we apply the same $(y^{\mathsf{SP}},\ldots,y^{\mathsf{SP}})$ to the deterministic problem (**D**) and accordingly have $n^{\mathsf{D}'}$ and the deterministic revenue $V^{\mathsf{D}'}(\alpha,T)$.

The proof of Corollary B.2 follows exactly the same logic of the proof of Lemma B.5. Whenever we use Lemma B.4 in the proof of Lemma B.5, we replace these with Corollary B.1. Whenever we use Lemma III.3 to bound $\mathbb{E}\left|\mathbf{y}^{\mathsf{CE}}\left(\hat{N}_{t-1}^m,t\right)-y_t^{\mathsf{D}}\right|$, we do not need them because we have zero gap between two sequences of y, that is $\mathbb{E}\left|\mathbf{y}^{\mathsf{SP}}\left(N_{t-1}^m,t\right)-y^{\mathsf{SP}}\right|=0$. □

B.3.6 Proof of Lemma B.7

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 $\mathbf{y}^{\mathsf{SP}}$ with $N_0^m=\alpha m$. Define $I_t^m\triangleq N_t^m/m$.

First we notice that

$$V^{\alpha,\mathbf{y}^{\mathsf{SP}}}(m,T)\leq m\mathbb{E}\left(\sum_{t=1}^{T}x^{-1}(y^{\mathsf{SP}})\lambda\left(I_{t-1}^m,\alpha\right)y^{\mathsf{SP}}\right) \tag{B.11}$$

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

According to Corollary B.2 (see Appendix B.3.5), we know

$$mV^{\mathsf{SP}}(\alpha,T)-\mathcal{O}(\sqrt{m})\leq m\mathbb{E}\left(\sum_{t=1}^{T}x^{-1}(\mathbf{y}^{\mathsf{SP}})\lambda\left(I_{t-1}^m,\alpha\right)\mathbf{y}^{\mathsf{SP}}\right)\leq mV^{\mathsf{SP}}(\alpha,T)+\mathcal{O}(\sqrt{m}). \tag{B.12}$$

Plugging (B.12) into RHS of (B.11), we get

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

□

BIBLIOGRAPHY

BIBLIOGRAPHY

Aït-Sahalia, Y. (2002), Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach, *Econometrica*, *70*(1), 223–262.

Alonso, R., and O. Câmara (2016a), Persuading voters, *American Economic Review*, *106*(11), 3590–3605.

Alonso, R., and O. Câmara (2016b), Political disagreement and information in elections, *Games and Economic Behavior*, *100*, 390–412.

Ambrus, A., and S. E. Lu (2010), Robust almost fully revealing equilibria in multi-sender cheap talk, *Tech. rep.*, mimeo.

Araman, V. F., and R. Caldentey (2009), Dynamic pricing for nonperishable products with demand learning, *Operations research*, *57*(5), 1169–1188.

Araman, V. F., and R. Caldentey (2010), Revenue management with incomplete demand information, in *Wiley Encyclopedia of Operations Research and Management Science*, edited by J. Cochran, John Wiley & Sons, Inc.

Arieli, I., and Y. Babichenko (2019), Private Bayesian persuasion, *Journal of Economic Theory*, *182*, 185–217.

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

Banker, R. D., I. Khosla, and K. K. Sinha (1998), Quality and competition, *Management Science*, *44*(9), 1179–1192.

Baptista, R. (1999), The diffusion of process innovations: A selective review, *International Journal of the Economics of Business*, *6*(1), 107–129.

Bardenet, R., O.-A. Maillard, et al. (2015), Concentration inequalities for sampling without replacement, *Bernoulli*, *21*(3), 1361–1385.

Bass, F. M. (1969), A new product growth for model consumer durables, *Management science*, *15*(5), 215–227.

Bass, F. M., and A. V. Bultez (1982), A note on optimal strategic pricing of technological innovations, *Marketing Science*, *1*(4), 371–378.

Bass, F. M., T. V. Krishnan, and D. C. Jain (1994), Why the Bass model fits without decision variables, *Marketing science*, *13*(3), 203–223.

Battaglini, M. (2002), Multiple referrals and multidimensional cheap talk, *Econometrica*, *70*(4), 1379–1401.

Bayus, B. L. (1992), The dynamic pricing of next generation consumer durables, *Market-*

ing Science, *11*(3), 251–265.

BBC, D. F. (2018), The hype machine: Streetware and the business of scarcity.

Bertsimas, D., and G. Perakis (2006), Dynamic pricing: A learning approach, in *Mathematical and Computational Models for Congestion Charging*, edited by S. Lawphongpanich, D. W. Hearn, and M. J. Smith, pp. 45–79, Springer US, Boston, MA.

Besbes, O., and D. Sauré (2014), Dynamic pricing strategies in the presence of demand shifts, *Manufacturing & Service Operations Management*, *16*(4), 513–528.

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

Besbes, O., and A. Zeevi (2011), On the minimax complexity of pricing in a changing environment, *Operations research*, *59*(1), 66–79.

Besbes, O., Y. Gur, and A. Zeevi (2015), Non-stationary stochastic optimization, *Operations research*, *63*(5), 1227–1244.

Bickel, P., D. Choi, X. Chang, H. Zhang, et al. (2013), Asymptotic normality of maximum likelihood and its variational approximation for stochastic blockmodels, *The Annals of Statistics*, *41*(4), 1922–1943.

Bitran, G. R., and S. V. Mondschein (1997), Periodic pricing of seasonal products in retailing, *Management Science*, *43*(1), 64–79.

Böker, F. (1987), A stochastic first purchase diffusion model: A counting process approach, *Journal of Marketing Research*, pp. 64–73.

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

Bradley, R. A., and J. J. Gart (1962), The asymptotic properties of ML estimators when sampling from associated populations, *Biometrika*, *49*(1/2), 205–214.

Brémaud, P. (1981), *Point processes and queues: martingale dynamics*, vol. 50, Springer.

Broder, J., and P. Rusmevichientong (2012), Dynamic pricing under a general parametric choice model, *Operations Research*, *60*(4), 965–980.

Brown, L. (2020), *Top 10 Richest YouTubers in 2020*.

Bumpensanti, P., and H. Wang (2020), A re-solving heuristic with uniformly bounded loss for network revenue management, *Management Science*.

Business Wire, K. H. (2019), Apple reports first quarter results: Services, Mac and Wearables set new all-time revenue records, EPS reaches all-time high at $4.18.

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

Candogan, O. (2019), Persuasion in networks: Public signals and k-cores, in *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 133–134.

Candogan, O., and K. Drakopoulos (2020), Optimal signaling of content accuracy: Engagement vs. misinformation, *Operations Research*.

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

Chakraborty, A., and R. Harbaugh (2014), Persuasive puffery, *Marketing Science*, *33*(3), 382–400.

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

Chen, Y. (2011), Perturbed communication games with honest senders and naive receivers, *Journal of Economic Theory*, *146*(2), 401–424.

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

Chen, Y., N. Kartik, and J. Sobel (2008), Selecting cheap-talk equilibria, *Econometrica*, *76*(1), 117–136.

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

Chow, G. C. (1960), Tests of equality between sets of coefficients in two linear regressions, *Econometrica: Journal of the Econometric Society*, pp. 591–605.

Cohen, A., and E. Solan (2013), Bandit problems with lévy processes, *Mathematics of Operations Research*, *38*(1), 92–107.

Cover, T. M. (1999), *Elements of information theory*, John Wiley & Sons.

Crawford, V. P., and J. Sobel (1982), Strategic information transmission, *Econometrica*, pp. 1431–1451.

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

Dellarocas, C. (2006), Strategic manipulation of internet opinion forums: Implications for consumers and firms, *Management Science*, *52*(10), 1577–1593.

den Boer, A., and B. Zwart (2015), Dynamic pricing and learning with finite inventories, *Operations research*, *63*(4), 965–978.

den Boer, A. V. (2015a), Dynamic pricing and learning: Historical origins, current research, and new directions, *Surveys in operations research and management science*, *20*(1), 1–18.

den Boer, A. V. (2015b), Tracking the market: Dynamic pricing and learning in a changing environment, *European journal of operational research*, *247*(3), 914–927.

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

Dolan, R. J., and A. P. Jeuland (1981), Experience curves and dynamic demand models: Implications for optimal pricing strategies, *The Journal of Marketing*, pp. 52–62.

Duffie, D., and P. Glynn (2004), Estimation of continuous-time markov processes sampled at random time intervals, *Econometrica*, *72*(6), 1773–1808.

Duffie, D., and K. J. Singleton (1993), Simulated moments estimation of markov models of asset prices, *Econometrica*, *61*(4), 929–952.

Eren, S. S., and C. Maglaras (2010), Monopoly pricing with limited demand information, *Journal of revenue and pricing management*, *9*(1-2), 23–48.

Ethier, S. N., and T. G. Kurtz (2005), *Markov processes: Characterization and convergence*, John Wiley & Sons.

Farias, V. F., and B. Van Roy (2010), Dynamic pricing with a prior on market response, *Operations Research*, *58*(1), 16–29.

Farrell, J. (1993), Meaning and credibility in cheap-talk games, *Games and Economic Behavior*, *5*(4), 514–531.

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

Ferreira, K. J., B. H. A. Lee, and D. Simchi-Levi (2015), Analytics for an online retailer: Demand forecasting and price optimization, *Manufacturing & Service Operations Management*, *18*(1), 69–88.

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

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

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

Gilligan, T. W., and K. Krehbiel (1989), Asymmetric information and legislative rules with a heterogeneous committee, *American Journal of Political Science*, *33*(2), 459–490.

Godes, D., and D. Mayzlin (2009), Firm-created word-of-mouth communication: Evidence from a field test, *Marketing Science*, *28*(4), 721–739.

Gordon, S. (2011), Iteratively stable cheap talk, *Department of Economics, Ecole des Sciences Politiques, Paris.*

Guseo, R., and C. Mortarino (2013), Within-brand and cross-brand word-of-mouth for sequential multi-innovation diffusions, *IMA Journal of Management Mathematics*, *25*(3), 287–311.

Hamami, T. (2019), Network effects, bargaining power, and product review bias: Theory and evidence, *The Journal of Industrial Economics*, *67*(2), 372–407.

Harrison, J. M., N. B. Keskin, and A. Zeevi (2012), Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution, *Management Science*, *58*(3), 570–586.

Hoadley, B. (1971), Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case, *The Annals of mathematical statistics*, *42*(6), 1977–1991.

Horsky, D. (1990), A diffusion model incorporating product benefits, price, income and information, *Marketing Science*, *9*(4), 342–365.

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

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

Kahn, K. B. (2006), *New Product Forecasting: An applied approach*, M.E. Sharpe, Armonk, NY.

Kalish, S. (1983), Monopolist pricing with dynamic demand and production cost, *Marketing Science*, *2*(2), 135–159.

Kalyanasundaram, S., E. K. Chong, and N. B. Shroff (2002), Markov decision processes with uncertain transition rates: Sensitivity and robust control, in *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, vol. 4, pp. 3799–3804, IEEE.

Kamenica, E., and M. Gentzkow (2011), Bayesian persuasion, *American Economic Review*, *101*(6), 2590–2615.

Kamrad, B., S. S. Lele, A. Siddique, and R. J. Thomas (2005), Innovation diffusion uncertainty, advertising and pricing policies, *European Journal of Operational Research*, *164*(3), 829–850.

Kartik, N. (2009), Strategic communication with lying costs, *The Review of Economic Studies*, *76*(4), 1359–1395.

Keiding, N., et al. (1975), Maximum likelihood estimation in the birth-and-death process, *The Annals of Statistics*, *3*(2), 363–372.

Keskin, N. B., and A. Zeevi (2014), Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies, *Operations Research*, *62*(5), 1142–1167.

Keskin, N. B., and A. Zeevi (2016), Chasing demand: Learning and earning in a changing environment, *Mathematics of Operations Research*, *42*(2), 277–307.

Kessler, M. (1995), Martingale estimating functions for a markov chain, *Working paper.*

Kessler, M. (1997), Estimation of an ergodic diffusion from discrete observations, *Scandinavian Journal of Statistics*, *24*(2), 211–229.

Kessler, M. (2000), Simple and explicit estimating functions for a discretely observed diffusion process, *Scandinavian Journal of Statistics*, *27*(1), 65–82.

Klear (2019), *2019 State of Influencer Marketing Report.*

Kleywegt, A. J., A. Shapiro, and T. Homem-de Mello (2002), The sample average approximation method for stochastic discrete optimization, *SIAM Journal on Optimization*, *12*(2), 479–502.

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

Kozinets, R. V., K. De Valck, A. C. Wojnicki, and S. J. Wilner (2010), Networked narratives: Understanding word-of-mouth marketing in online communities, *Journal of Marketing*, *74*(2), 71–89.

Krishna, V. (2001), Asymmetric information and legislative rules: Some amendments, *American Political science review*, *95*(2), 435–452.

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

Krishnan, T. V., F. M. Bass, and D. C. Jain (1999), Optimal pricing strategy for new products, *Management Science*, *45*(12), 1650–1663.

Krishnan, T. V., F. M. Bass, and V. Kumar (2000), Impact of a late entrant on the diffusion of a new product/service, *Journal of marketing research*, *37*(2), 269–278.

Küchler, U., and M. Sorensen (2006), *Exponential families of stochastic processes*, Springer Science & Business Media.

Kurtz, T. G. (1971), Limit theorems for sequences of jump markov processes approximating ordinary differential processes, *Journal of Applied Probability*, *8*(2), 344–356.

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

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

Li, H. (2020), Optimal pricing under diffusion-choice models, *Operations Research.*

Li, H., and W. T. Huh (2012), Optimal pricing for a short life-cycle product when customer price-sensitivity varies over time, *Naval Research Logistics (NRL)*, *59*(7), 552–576.

Lim, A. E., and J. G. Shanthikumar (2007), Relative entropy, exponential utility, and robust dynamic pricing, *Operations Research*, *55*(2), 198–214.

Lin, K. Y. (2006), Dynamic pricing with real-time demand learning, *European Journal of Operational Research*, *174*(1), 522–538.

Lingenbrink, D., and K. Iyer (2018), Signaling in online retail: Efficacy of public signals, in *Proceedings of the 13th Workshop on Economics of Networks, Systems and Computation*, pp. 1–1.

Lingenbrink, D., and K. Iyer (2019), Optimal signaling mechanisms in unobservable queues, *Operations Research*, *67*(5), 1397–1416.

Lorentz, G. G. (2013), *Bernstein polynomials*, American Mathematical Soc.

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

Mahajan, V., E. Muller, and F. M. Bass (1995), Diffusion of new products: Empirical generalizations and managerial uses, *Marketing science*, *14*(3_supplement), G79–G88.

Mayzlin, D. (2006), Promotional chat on the internet, *Marketing science*, *25*(2), 155–163.

Mayzlin, D., Y. Dover, and J. Chevalier (2014), Promotional reviews: An empirical investigation of online review manipulation, *American Economic Review*, *104*(8), 2421–55.

MediaKix (2019), *Influencer Marketing 2019: Key Statistics From Our Influencer Marketing Survey.*

Nilim, A., and L. El Ghaoui (2004), Robust markov decision processes with uncertain transition matrices, Ph.D. thesis, University of California, Berkeley.

Nilim, A., and L. El Ghaoui (2005), Robust control of markov decision processes with uncertain transition matrices, *Operations Research*, *53*(5), 780–798.

Niu, S.-C. (2002), A stochastic formulation of the Bass model of new-product diffusion, *Mathematical problems in Engineering*, *8*(3), 249–263.

Niu, S.-C. (2006), A piecewise-diffusion model of new-product demands, *Operations Research*, *54*(4), 678–695.

Norman, M. F., et al. (1974), A central limit theorem for markov processes that move by small steps, *The Annals of Probability*, *2*(6), 1065–1074.

Norton, J. A., and F. M. Bass (1987), A diffusion theory model of adoption and substitution for successive generations of high-technology products, *Management science*, *33*(9), 1069–1086.

Qi, A., H.-S. Ahn, and A. Sinha (2017), Capacity investment with demand learning, *Operations Research*, *65*(1), 145–164.

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

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

Rayo, L., and I. Segal (2010), Optimal information disclosure, *Journal of Political Economy*, *118*(5), 949–987.

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

Robinson, B., and C. Lakhani (1975), Dynamic price models for new-product planning, *Management science*, *21*(10), 1113–1122.

Roussas, G. G. (1969), Nonparametric estimation in Markov processes, *Annals of the Institute of Statistical Mathematics*, *21*(1), 73–87.

Sapra, A., V.-A. Truong, and R. Q. Zhang (2010), How much demand should be fulfilled?, *Operations Research*, *58*(3), 719–733.

Savin, S., and C. Terwiesch (2005), Optimal product launch times in a duopoly: Balancing life-cycle revenues with product cost, *Operations Research*, *53*(1), 26–47.

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

Schmittlein, D. C., and V. Mahajan (1982), Maximum likelihood estimation for an innovation diffusion model of new product acceptance, *Marketing science*, *1*(1), 57–78.

Serfling, R. J. (1974), Probability inequalities for the sum in sampling without replacement, *The Annals of Statistics*, pp. 39–48.

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

Shen, W., I. Duenyas, and R. Kapuscinski (2013), Optimal pricing, production, and inventory for new product diffusion under supply constraints, *Manufacturing & Service Operations Management*, *16*(1), 28–45.

Shreve, S. E. (2004), *Stochastic calculus for finance II: Continuous-time models*, vol. 11, Springer Science & Business Media.

Smith, S. A., and D. D. Achabal (1998), Clearance pricing and inventory policies for retail chains, *Management Science*, *44*(3), 285–300.

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

Sobel, J. (2013), Giving and receiving advice, *Advances in Economics and Econometrics*, *1*, 305–341.

Srinivasan, V., and C. H. Mason (1986), Technical note—Nonlinear least squares estimation of new product diffusion models, *Marketing science*, *5*(2), 169–178.

Stonemaier Games (2013), Kickstarter lesson #62: Early bird pledge levels.

Taneja, I. J., and P. Kumar (2004), Relative information of type s, csiszár's f-divergence, and information inequalities, *Information Sciences*, *166*(1-4), 105–125.

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

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

USA TODAY, J. G. (2019), Did Apple retail prices get too high in 2018? Consumers say yes.

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

Vardhman, R. (2019), *Under the Influence – 80+ Influencer Marketing Statistics*.

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

Wang, Z., S. Deng, and Y. Ye (2014), Close the gaps: A learning-while-doing algorithm for single-product revenue management problems, *Operations Research*, *62*(2), 318–331.

Whittle, P. (1988), Restless bandits: Activity allocation in a changing world, *Journal of applied probability*, *25*(A), 287–298.

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

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

ZDNet, A. K.-H. (2019), Apple CEO on iPhone sales drop: 'Yes, i do think that price is a factor'.

Zhao, M., and M. Xie (1996), On maximum likelihood estimation for a general nonhomogeneous poisson process, *Scandinavian journal of statistics*, pp. 597–607.

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