Data-Driven Pricing for a New Product

Mengzhenyu Zhang, Hyun-Soo Ahn, Joline Uichanco
Stephen M. Ross School of Business, University of Michigan, Ann Arbor, MI, USA
zhenyuzh, hsahn, jolineu@umich.edu

Decisions regarding new products are often difficult to make, and mistakes can have grave consequences for a firm's bottom line. Often, firms lack important information about a new product such as its potential market size and the speed of its adoption by consumers. One of the most popular frameworks that has been used for modeling new product adoption is the Bass model (Bass 1969). While the Bass model and its many variants have been used to study dynamic pricing of new products, the vast majority of these models require *a priori* knowledge of parameters that can only be estimated from historical data or guessed using institutional knowledge. In this paper, 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 selling horizon. We extend the generalized Bass model to a stochastic setting by modeling adoption through a continuous-time Markov chain where the adoption rate depends on the selling price and on the number of past sales. We study a pricing problem where the parameters of this demand model are unknown, but the seller can utilize real-time demand data for learning the parameters. We propose two simple and computationally tractable pricing policies with $\mathcal{O}(\ln m)$ regret where m is the market size.

Key words: Bass adoption model, Data-driven pricing, Demand learning

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 (e.g., 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. Lenk and Rao (1990) propose a Hierarchical Bayes procedure to gain information from different products which share some common structures and to use this information for forecasting sales of a new product. 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 revenues 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 paper, we study the interplay between demand learning and dynamic pricing for a new product. 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 contributions of this paper are 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, we establish the convergence properties of the estimation error of the ML estimators. The challenge in proving the latter result is that 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.
- *Performance guarantee for tractable pricing policies.* 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 regret that is in the order of the log of the market size. We prove a fundamental lower bound on the regret of any pricing-and-learning policy, and show that the regret of our policies match this limit.

1.1. Review of related literature

Our paper 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.

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.

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.

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. 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.

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

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.

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:

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

We can compute its solution as

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

[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 examples.

The generalized Bass model (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 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 has purchased 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

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

where

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

Note that $\xi(j)$ is the portion of the adoption rate unaffected by price. From (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,

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

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

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

Hence, the Markovian Bass model captures the demand dynamics in the generalized Bass model (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 on 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 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 converges to the deterministic Bass curve.

Proposition 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 for any $t \geq 0$:*

$$\frac{D^{r,m_0}_t}{m_0} \to F^r_t \textit{ almost surely as } m_0 \to \infty, \tag{2.7}$$

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

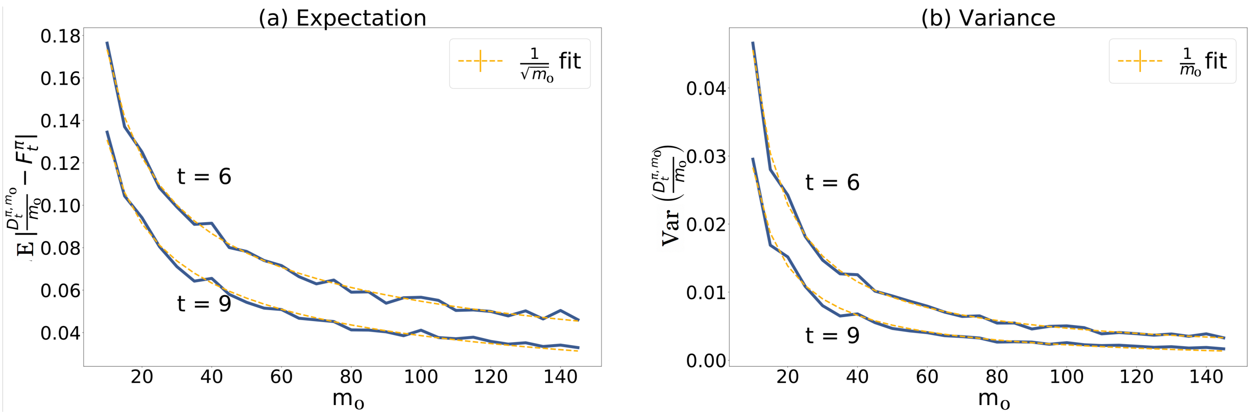


Figure 1 Convergence of expectation and variance of $D_t^{\pi,m_0}/m_0$ as m_0 increases.

Proposition 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 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 in an asymptotic regime.

Lemma 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*

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

Figure 1 illustrates Lemma 2 and Proposition 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 $\frac{1}{m_0}$.

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 can use the data to infer the unknown demand parameters. Without knowing θ_0, the seller needs to choose a pricing policy π that is $\mathcal{F}_t$-adapted.

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. 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.

To evaluate a pricing policy π, we will use a performance measure called the *regret*:

$$\text{Regret}(\pi) := R^* - R(\pi) := R^* - \mathbb{E}_{\theta_0}\left[\int_0^T r_t^\pi \mathrm{d}D_t^\pi\right] \tag{2.9}$$

where R^* is the optimal expected cumulative revenue if the seller knows the true value of θ_0, and $R(\pi)$ is the expected cumulative revenue of the pricing policy π. The expectation is taken with respect to the Markovian Bass model with parameter vector θ_0.

The Bass diffusion model has a long tradition in the marketing literature of being used to derive 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.

Throughout the paper, we assume that the marketing effort function $x(\cdot)$ satisfies certain regularity properties.

Assumption 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 $|x'(r)|\le M$, for all $-\infty<r<\infty$;

ii. (Non-negativity.) There exist non-negative constants $\bar{x}^u,\bar{x}^l$, such that $\bar{x}^l t\le\int_0^t x(r_s)\mathrm{d}s\le\bar{x}^u t$, for all $r=(r_s,s\ge 0)$ where $-\infty<r_s<\infty$ for all $s\ge 0$;

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 $2x'(r)^2-x(r)x''(r)\ge C_d>0$;

v. (Boundedness of revenue.) There exist constants C_x, C_{xx} such that, for $f(r):=rx(r)$, $|f(r)|\le C_x$, and $|f''(r)|\le C_{xx}$ for all $-\infty<r<\infty$.

Assumption 1(i)–(iii) are innocuous as they guarantee that the market effort function is decreasing in price and is sufficiently smooth. Assumption 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 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 will propose pricing policies when the Markovian Bass model has an unknown parameter vector θ_0. To analyze their regret, we first should study the optimal expected revenue when θ_0 is known. This is the objective of the next section.

3. Optimal pricing policy with complete information

If the firm knows the demand model parameter vector θ_0, it can maximize its expected revenue by solving the following optimal control problem:

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

where the optimal expected revenue is R^*. Here, the equality follows from the tower property of conditional expectation.

To solve the optimal control problem (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 (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 (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 EC.1(iii).

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

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

where λ is the adoption rate defined in (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 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* (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*

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

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

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

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 (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 (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 (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

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

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 (3.4) for the deterministic Bass model to the condition (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 (3.2) and (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 (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 (3.2) and rearranging terms, we have

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

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 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 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 (3.3) numerically using finite differences, a conventional technique for solving partial differential equations numerically. We describe this method in the e-companion (Appendix EC.1).

4. Data-driven dynamic pricing with unknown parameters

In this section, we propose pricing policies under 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

[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.

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 3 later establishes that when wrong model parameters are initially inferred and the data is not used to correct the wrong inference, the regret (relative to R^*) can grow at least linearly in the true market size m_0. This motivates our pricing policies that learn the unknown parameters from the price and sales data.

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

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

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

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

This results in the following log-likelihood function

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

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 $\max_{\theta \geq 0} \mathcal{L}_t(\widehat{\mathbf{U}}_t \mid \theta)$. It is difficult to show the joint concavity of the log-likelihood function in θ. Hence, we perform the following variable transformation:

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

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:

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

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

Proposition 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 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 (4.4). (Transforming (4.2) using (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 as the sample size increases 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

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

not identically distributed samples. 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 3. *For any fixed time $t > 0$ and $k \geq 3$,*

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

for some $\alpha_\theta > 0$ that is independent of m_0, t and k.

The variance of $\hat{q}_t$ grows as D_t^π / m_0 approaches zero. However, as we have shown in the proof of Lemma 3, the bound on the variance of $\hat{q}_t$ does not depend on m_0, t or k if $\hat{q}_t$ is estimated from the ML model under the simple parameter transformation $(p - q, q, m)$. The Lemma 3 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.

4.2. Data-driven pricing policies

We next develop sensible data-driven dynamic pricing policies that (1) utilize data in a computationally efficient way, and (2) have performance guarantees on regret. The policies utilize the Markovian Bass price function introduced in Section 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 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.

4.2.1. Pricing policy with continuous price changes. We first propose a policy referred to as MBP-MLE (outlined in Algorithm 1). At time t, this policy offers the Markovian Bass price of Section 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 regret of MBP-MLE grows sublinearly in the market size at the rate of $\mathcal{O}(\ln m_0)$ (Theorem 4).

The policy starts with an initial price $r_t = 0$ for $t \in [0, t_3]$, where t_3 is the time of the third adoption. When $t \in (t_3, T]$, the seller uses the accrued data to solve for the ML solution $\hat{\theta}_t$ and set the price at $r_t = r_t^*(\hat{\theta}_t, \widehat{D}_t)$. Since the ML model is only identifiable after time t_3 (see Proposition 2), the low initial price is chosen to lengthen the period, $(t_3, T]$, during which the policy uses the ML solution for pricing. However, any arbitrary initial price can be used.

Algorithm 1 MBP-MLE algorithm

Require: 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**
  if $\widehat{D}_s < 3$ **then**
    $r_s \leftarrow 0$ ▷ *Set price*
  else
    $\hat{\theta}_s \leftarrow \arg\max_{\theta} \ \mathcal{L}_t\left(\widehat{\mathbf{U}}_s \mid \theta\right)$ ▷ *Estimate parameter*
    $r_s \leftarrow \inf\left\{r : r \geq -\frac{x(r)}{x'(r)} - V(\widehat{D}_s + 1, T - \delta s; \hat{\theta}_s) + V(\widehat{D}_s, T - \delta s; \hat{\theta}_s)\right\}$ ▷ *Set price*
  $\widehat{\mathbf{U}}_s \leftarrow \widehat{\mathbf{U}}_{s-1} \cup \{(r_s, \widehat{D}_s + a_s)\}$, where a_s is the new sales in $[\delta s, \delta(s+1))$
  $s \leftarrow s + 1$ ▷ *Proceed to next period*

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 5.3, we describe how K and C can be chosen, even without knowing m_0, so that the regret of MBP-MLE-Limited is $\mathcal{O}(\ln m_0)$ (Theorem 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

Figure 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.

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 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 $j = 1, \ldots, C_i$), the adoption rate changes to $\lambda(d+j-1, r_j; \hat{\theta})$ where λ is defined in (2.3). Hence, the expected inter-adoption time between the $(d+j-1)^{\text{th}}$ and $(d+j)^{\text{th}}$ adoption is

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

This then determines the time of the next adoption $d+j$ under a deterministic model, assuming that the previous inter-adoption times $\Delta t_1, \Delta t_2, \ldots, \Delta t_{j-1}$ 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 $\{r_1, \ldots, r_{C_i}\}$ 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 (4.6), $r^{(i)}$ is the solution to

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

For the first price cycle ($i=0$), we assume that the policy starts with an initial price $r^{(0)}=0$. We also assume that $C_0 \geq 3$ so that there exists an ML estimator when the first price change is calculated. Similar to our reasoning with MBP-MLE, this low initial price is chosen to hasten the time t_{C_0} that the policy can start using the MLE solution for pricing. Algorithm 2 provides the outline for the MBP-MLE-Limited algorithm.

Algorithm 2 MBP-MLE-Limited algorithm

Require: Max horizon length T, subperiod length δ, 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 $d+j$*
    $t \leftarrow t - \Delta t_j$
  $\bar{r} \leftarrow \inf\left\{r : r \cdot x(r) \geq \frac{\sum_{j=1}^{C} r_j}{\sum_{j=1}^{C} 1/x(r_j)}\right\}$ ▷ *Calculate the price for the C adoptions*
  return $\bar{r}$
end function

$r_0 \leftarrow 0,\ s \leftarrow 1,\ i \leftarrow 1,\ \widehat{D}_0 \leftarrow 0,\ \widehat{\mathbf{U}}_0 \leftarrow \emptyset$ ▷ *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*
  $s \leftarrow s+1$ ▷ *Proceed to next period*

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 their regret, defined in (2.9). Specifically, we will derive asymptotic bounds on the regret of our proposed policies as the market size m_0 grows. We will establish that the regret of our proposed policies grow at most sublinearly with m_0 at the rate $\mathcal{O}(\ln m_0)$.

The challenge in bounding the regret 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 regret, we then need to establish a non-stationary relationship among regret, pricing errors and estimation errors.

Let $D^* = (D_t^*,\ t \ge 0)$ and $r^* = (r_t^*,\ t \ge 0)$ 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 1.2. Particularly, we are interested in the limiting behavior as m_0 goes to infinity.

Proposition 3. *If $\pi \in \Pi$ and $\{r_t^\pi, t \in [0,T]\}$ does not scale up with m_0, then*

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

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

This important proposition establishes that the regret 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 regret.

The idea behind the proof of Proposition 3 is that we can decompose the regret into two parts:

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

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 regret 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 regret due to the compounded word-of-mouth effects. We prove that, under Assumption 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 EC.2), the second part is dominated by the first part. The full proof can be found in the appendix.

The implication of Proposition 3 is that, to bound the regret 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 4. *Consider a parameter sequence $\{\bar{\theta}_t = (\bar{p}_t, \bar{q}_t, \bar{m}_t), t \geq 0\}$ where $\bar{\theta}_t$ is an $\mathcal{F}_t$-measurable random vector, and $\bar{p}_t, \bar{q}_t, \bar{m}_t$ are finite, $\bar{p}_t + \bar{q}_t > 0$ and $\bar{m}_t > 0$ for all $t \geq 0$ almost surely. If π is the policy that offers the Markovian Bass price with parameter $\bar{\theta}_t$, i.e., $r_t^\pi = r_t^*(\bar{\theta}_t, D_t^\pi)$, then for any $t \in (0, T]$,*

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

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 4 and Proposition 3 together enable us to bound the regret by the parameter estimation error. Hence, the regret 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 regret of data-driven pricing policies. This is usually accomplished by constructing a special case of the problem that satisfies Assumption 1, and showing that for any data-driven pricing policy, the worst-case regret 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 1.) Hence, the fundamental limit $\Omega(\ln m_0)$ serves as a benchmark for the regret of data-driven pricing policies in our problem setting.

THEOREM 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* $\mathrm{Regret}(\pi) = \Omega(\ln m_0)$.

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

In the remainder of this section, we will analyze the regret 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 1. In fact, we will show that the regret of our proposed policies are $\mathcal{O}(\ln m_0)$.

5.1. Regret 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 3 below states that the regret 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 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 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 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 2 and if $\mathcal{E}^2$ is the total estimation error such that $(\hat{p}_0 - p_0)^2/{p_0}^2 + (\hat{q}_0 - q_0)^2/{q_0}^2 + (\hat{m}_0 - m_0)^2/{m_0}^2 = \mathcal{E}^2$, then* $\mathrm{Regret}(\pi^s) = \Omega(\mathcal{E}^2 m_0)$.

5.2. Regret of MBP-MLE

We next establish an upper bound on the regret 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 3 and Lemma 4 is that the regret 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 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 4. *If π^{M} is the MBP-MLE policy, then* $\text{Regret}(\pi^{\mathsf{M}}) = \mathcal{O}(\ln m_0)$.

Note that Lemma 3, Lemma 4, and Proposition 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 3), which incidentally is also the weight applied to the pricing error in (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 regret of MBP-MLE, $\mathcal{O}(\ln m_0)$, matches the fundamental lower bound on the regret of any data-driven pricing policy (Theorem 2). In Section 5.5, we will discuss why learning appears to occur for free under MBP-MLE.

5.3. Regret of MBP-MLE-Limited

We now derive a performance bound for the MBP-MLE-Limited policy $\pi^{\text{M-Lim}}$, 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 $\sum_{i=1}^{K} C_i = \Theta(m_0)$ and that $C_0 = \Theta(1)$.

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

THEOREM 5. *Let $\pi^{\text{M-Lim}}$ 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,*

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

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 regret grows at most in logarithmic order. To see this, note that if $C_i = C_0 a^i$ for all $i = 1, \ldots, K$ for some base $a > 0$, then $\text{Regret}\left(\pi^{\text{M-Lim}}\right) = \mathcal{O}\left((1 + a) \ln m_0\right)$. Hence, the regret 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 regret 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 regret 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 $C_1 = \ldots = C_K = \Theta(m_0)$, then the regret 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 $(K+1)C_0 \geq \Theta(m_0)$.

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 3) in proving the bound on regret. 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.

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 $\gamma_0 = (1, \gamma_{0,1}, \gamma_{0,2} \dots, \gamma_{0,n})$. 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 (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:

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

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

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

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

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

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

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

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

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.

EXAMPLE 1. If $n=1$ then

$$\mathbf{Y}=\begin{pmatrix} 1-\widehat{r}_{t_1} & 1-\widehat{r}_{t_2} & 1-\widehat{r}_{t_3} & 1-\widehat{r}_{t_4} & 1-\widehat{r}_{t_5} & 1-\widehat{r}_{t_6} \\ 0 & 1-\widehat{r}_{t_2} & 2(1-\widehat{r}_{t_3}) & 3(1-\widehat{r}_{t_4}) & 4(1-\widehat{r}_{t_5}) & 5(1-\widehat{r}_{t_6}) \\ 0 & 1-\widehat{r}_{t_2} & 4(1-\widehat{r}_{t_3}) & 9(1-\widehat{r}_{t_4}) & 16(1-\widehat{r}_{t_5}) & 25(1-\widehat{r}_{t_6}) \\ \widehat{r}_{t_1} & \widehat{r}_{t_2} & \widehat{r}_{t_3} & \widehat{r}_{t_4} & \widehat{r}_{t_5} & \widehat{r}_{t_6} \\ 0 & \widehat{r}_{t_2} & 2\widehat{r}_{t_3} & 3\widehat{r}_{t_4} & 4\widehat{r}_{t_5} & 5\widehat{r}_{t_6} \\ 0 & \widehat{r}_{t_2} & 4\widehat{r}_{t_3} & 9\widehat{r}_{t_4} & 16\widehat{r}_{t_5} & 25\widehat{r}_{t_6} \end{pmatrix}.$$

If $\widehat{r}_{t_1}=\ldots=\widehat{r}_{t_6}$, then $\mathbf{Y}$ is *not* full rank. However, prices do not have to be all distinct for $\mathbf{Y}$ to be full rank. The price sequence, $\widehat{r}_{t_1}=0.8$ and $\widehat{r}_{t_2}=\widehat{r}_{t_3}=\ldots=\widehat{r}_{t_6}=0.9$ is an example. In fact, many initial price sequences result in a full rank matrix. For general n, since the Bernstein basis functions are known, the initial price sequence that result in a full rank $\mathbf{Y}$ matrix can be determined off-line (i.e., before time 0).

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:

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

We denote the solution to (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 (5.6) has a strictly concave objective function (under the Proposition 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 (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,γ_0). In fact, we can establish the rate of convergence of ML estimators, similar to Lemma 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 3 apply to ML estimators of $\mu_0=(\beta_0,\gamma_0)$. This gives us the following Lemma 3′.

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

$$\mathbb{E}_{\mu_0}\left(\left(\frac{\hat{p}_t - p_0}{p_0}\right)^2 + \left(\frac{\hat{q}_t - q_0}{q_0}\right)^2 + \left(\frac{\hat{m}_t - m_0}{m_0}\right)^2 \mid D_t^\pi = k\right) \leq \frac{\alpha_\theta}{k+1}, \quad \text{and}$$

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

where $\alpha_\theta, \alpha_\gamma$ are constants that are independent of m_0, t and k.

The next step is to establish the relationship between the pricing errors and the estimation errors, similar to Lemma 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 4′. *Consider a parameter sequence $\{\bar{\mu}_t = (\bar{p}_t, \bar{q}_t, \bar{m}_t, \bar{\gamma}_{t,0}, \ldots, \bar{\gamma}_{t,n}), t \geq 0\}$ where $\bar{\mu}_t$ is an $\mathcal{F}_t$-measurable random vector, and $\bar{p}_t, \bar{q}_t, \bar{m}_t, \bar{\gamma}_{t,j}$ are finite, and $\bar{p}_t + \bar{q}_t > 0$ and $\bar{m}_t > 0$ 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 $\bar{\mu}_t$, i.e., $r_t^\pi = r_t^*(\bar{\mu}_t, D_t^\pi)$, then for any $t \in (0, T]$,*

$$\mathbb{E}_{\mu_0}\left[|r_t^* - r_t^\pi|^2 \mid \mathcal{F}_t\right] = \Theta\left(\mathbb{E}_{\mu_0}\left[\left(\tfrac{\bar{p}_t - p_0}{p_0}\right)^2 + \left(\tfrac{\bar{q}_t - q_0}{q_0}\right)^2 + \left(\tfrac{\bar{m}_t - m_0}{m_0}\right)^2 \mid \mathcal{F}_t\right]\right) + \Theta\left(\mathbb{E}_{\mu_0}\left[\|\bar{\gamma}_t - \gamma_0\|^2 \mid \mathcal{F}_t\right]\right) + \mathcal{O}\left(\frac{1}{m_0}\right).$$

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

5.5. Discussion of why learning occurs for "free"

Recall that our bound $\mathcal{O}(\ln m_0)$ on MBP-MLE and MBP-MLE-Limited coincides with the fundamental lower bound on regret in Theorem 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 $\mathcal{O}(\ln m_0)$ regret 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 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

(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 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.

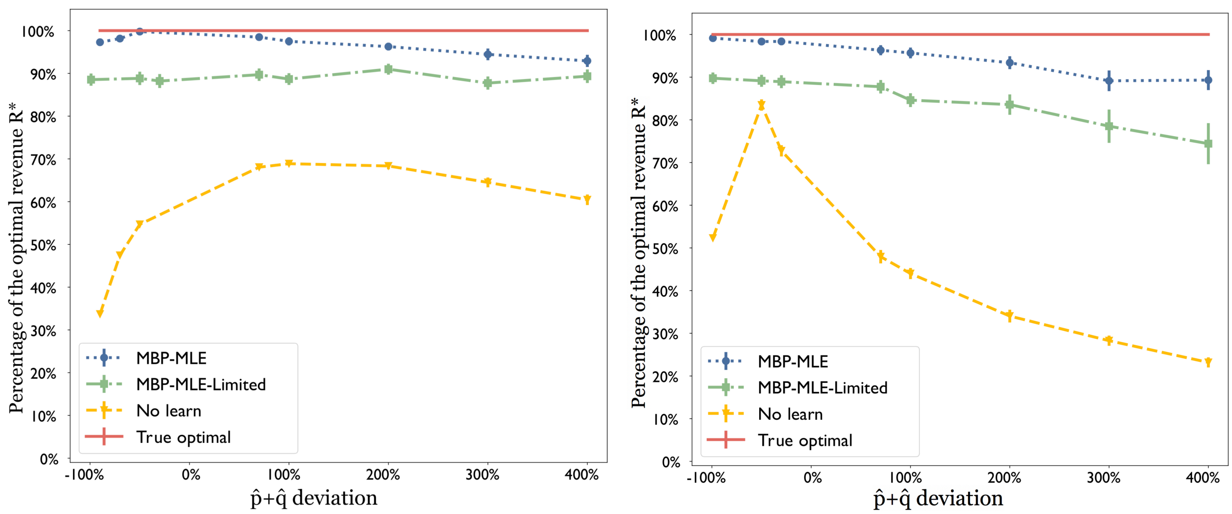
MBP-MLE allows continuous price changes, so the benefit from the increasingly accurate ML estimators is immediately realized through pricing that exploits the current estimates. This explains why the regret is $\mathcal{O}(\ln m_0)$ even without changing prices for the explicit purpose of price exploration. 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 3 and Lemma 3′), the changes in the estimators are only reflected onto the price at the limited price change epochs. Therefore, a key to limiting the regret 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 impact on the total revenue. One possibility is to set the adoptions between price changes to increase exponentially. Theorem 5 and the ensuing discussion show that such a choice achieves $\mathcal{O}(\ln m_0)$ regret. 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.

6. Numerical Study

In numerical studies, we compare the regret of MBP-MLE, MBP-MLE-Limited, and a no-learning pricing policy (i.e., MBP policy based on prior parameter values without updating). For MBP-MLE and the no-learning pricing policy, the initial prices are the MBP prices solved from an initial estimate $\hat{\theta} = (\hat{p}, \hat{q}, \hat{m})$. The prices of the two policies start deviating after three adoptions (i.e., after the MLE model is identifiable). The initial price for MBP-MLE-Limited (i.e., for the first C_0 customers) is the result of function LIMITED-PRICE$(\hat{\theta}, 3, 0, T)$ defined in Algorithm 2. The price change epochs are $C_i = 3 \cdot 2^i$ for $i = 0, 1, \ldots$

Figure 3 shows how the revenues are changing with respect to the initial modeling error (i.e., by how much $\hat{\theta}$ is different from θ_0). In these experiments, we assume that the true parameters are $\theta_0 = (p_0, q_0, m_0) = (0.4, 0.6, 100)$ and that $T = 40$. The horizontal axis is the percentage deviation δ_{pq} where $\hat{p} = (1 + \delta_{pq})p_0$ and $\hat{q} = (1 + \delta_{pq})q_0$. We vary the percentage deviation δ_{pq} from -90% to



(a) Deviation of $\hat{m} = -50\%$ (b) Deviation of $\hat{m} = 450\%$

Figure 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.

400% and plot the ratio $R(\pi)/R^*$ of the three policies. Figures 3 (a) and (b) display the plots for different deviations $\delta_m = -50\%$ and $\delta_m = 450\%$, where $\hat{m} = (1+\delta_m)m_0$.

While the revenue R^* from the oracle 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 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, MBP-MLE and MBP-MLE-Limited give 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.

An insight from Figure 3 is that the regret of MBP-MLE and of MBP-MLE-Limited are smallest when the initial estimate $\hat{\theta}$ is the smallest, which corresponds to a low initial price. Since the learning accuracy of MLE is directly associated with number of cumulative adopters, a high initial price leads to slower adoption and a larger regret for both MBP-MLE and MBP-MLE-Limited. Hence, the best numerical performance of our policies result when the initial estimate $\hat{\theta}$ is small.

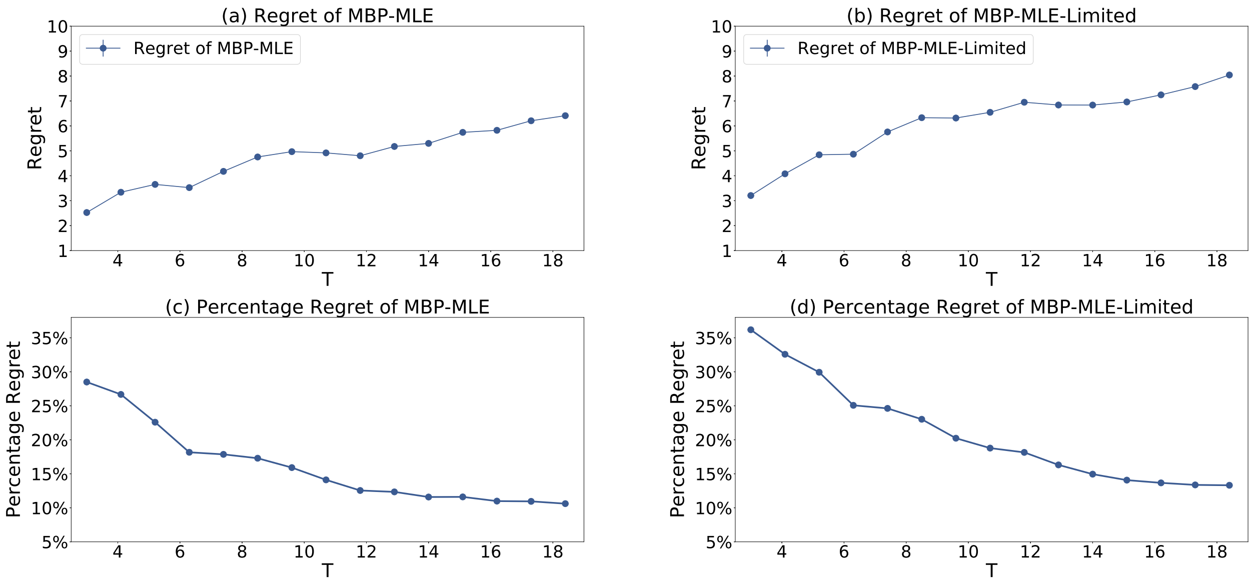


Figure 4 Cumulative regret and percentage cumulative regret of MBP-MLE and MBP-MLE-Limited.

Figures 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 regret of MBP-MLE and MBP-MLE-Limited, shown in Panels (c) and (d), decrease rapidly. Each dot in the figure is the average regret 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 4, we clearly see that the regret 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 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%.

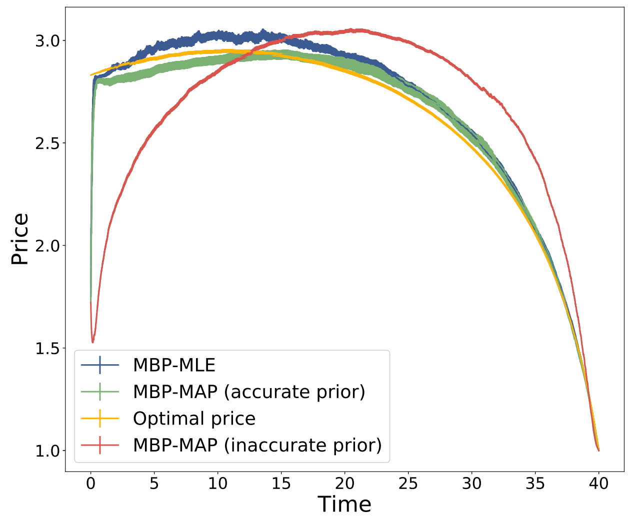


Figure 5 The true values of the parameters are $m_0 = 150, p_0 = 0.4, q_0 = 0.6, T = 40$. We run 100 experiments of the price sample paths and plot the average with the 95% confidence interval.

Figure 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.

7. Conclusion

This paper 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 of a new product, our paper 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.

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 regret 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 regret 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 regret 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.

Acknowledgements

The authors are grateful for the valuable comments and suggestions for improvements from the associate editor and three anonymous reviewers.

References

Aït-Sahalia Y (2002) Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica* 70(1):223–262.

Araman VF, Caldentey R (2009) Dynamic pricing for nonperishable products with demand learning. *Operations research* 57(5):1169–1188.

Araman VF, Caldentey R (2010) Revenue management with incomplete demand information. Cochran J, ed., *Wiley Encyclopedia of Operations Research and Management Science* (John Wiley & Sons, Inc.).

Baptista R (1999) The diffusion of process innovations: A selective review. *International Journal of the Economics of Business* 6(1):107–129.

Bardenet R, Maillard OA, et al. (2015) Concentration inequalities for sampling without replacement. *Bernoulli* 21(3):1361–1385.

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

Bass FM, Bultez AV (1982) A note on optimal strategic pricing of technological innovations. *Marketing Science* 1(4):371–378.

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

Bayus BL (1992) The dynamic pricing of next generation consumer durables. *Marketing Science* 11(3):251–265.

BBC DF (2018) The hype machine: Streetware and the business of scarcity. URL `http://www.bbc.com/capital/story/20180205-the-hype-machine-streetwear-and-the-business-of-scarcity`.

Bertsimas D, Perakis G (2006) Dynamic pricing: A learning approach. Lawphongpanich S, Hearn DW, Smith MJ, eds., *Mathematical and Computational Models for Congestion Charging*, 45–79 (Boston, MA: Springer US).

Besbes O, Gur Y, Zeevi A (2015) Non-stationary stochastic optimization. *Operations research* 63(5):1227–1244.

Besbes O, Sauré D (2014) Dynamic pricing strategies in the presence of demand shifts. *Manufacturing & Service Operations Management* 16(4):513–528.

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

Besbes O, Zeevi A (2011) On the minimax complexity of pricing in a changing environment. *Operations research* 59(1):66–79.

Bickel P, Choi D, Chang X, Zhang H, et al. (2013) Asymptotic normality of maximum likelihood and its variational approximation for stochastic blockmodels. *The Annals of Statistics* 41(4):1922–1943.

Böker F (1987) A stochastic first purchase diffusion model: A counting process approach. *Journal of Marketing Research* 64–73.

Bradley RA, Gart JJ (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*, volume 50 (Springer).

Broder J, Rusmevichientong P (2012) Dynamic pricing under a general parametric choice model. *Operations Research* 60(4):965–980.

Business Wire KH (2019) Apple reports first quarter results: Services, Mac and Wearables set new all-time revenue records, EPS reaches all-time high at $4.18. URL `https://www.businesswire.com/news/home/20190129005873/en/Apple-Reports-Quarter-Results`.

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

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

Chow GC (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society* 591–605.

Cover TM (1999) *Elements of information theory* (John Wiley & Sons).

den Boer A, Zwart B (2015) Dynamic pricing and learning with finite inventories. *Operations research* 63(4):965–978.

den Boer AV (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 AV (2015b) Tracking the market: Dynamic pricing and learning in a changing environment. *European journal of operational research* 247(3):914–927.

Dolan RJ, Jeuland AP (1981) Experience curves and dynamic demand models: Implications for optimal pricing strategies. *The Journal of Marketing* 52–62.

Duffie D, Glynn P (2004) Estimation of continuous-time markov processes sampled at random time intervals. *Econometrica* 72(6):1773–1808.

Duffie D, Singleton KJ (1993) Simulated moments estimation of markov models of asset prices. *Econometrica* 61(4):929–952, ISSN 00129682, 14680262, URL `http://www.jstor.org/stable/2951768`.

Eren SS, Maglaras C (2010) Monopoly pricing with limited demand information. *Journal of revenue and pricing management* 9(1-2):23–48.

Ethier SN, Kurtz TG (2005) *Markov processes: Characterization and convergence* (John Wiley & Sons).

Farias VF, Van Roy B (2010) Dynamic pricing with a prior on market response. *Operations Research* 58(1):16–29.

Ferreira KJ, Lee BHA, Simchi-Levi D (2015) Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management* 18(1):69–88.

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

Guseo R, Mortarino C (2013) Within-brand and cross-brand word-of-mouth for sequential multi-innovation diffusions. *IMA Journal of Management Mathematics* 25(3):287–311.

Harrison JM, Keskin NB, Zeevi A (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.

Kahn KB (2006) *New Product Forecasting: An applied approach* (Armonk, NY: M.E. Sharpe).

Kalish S (1983) Monopolist pricing with dynamic demand and production cost. *Marketing Science* 2(2):135–159.

Kalyanasundaram S, Chong EK, Shroff NB (2002) Markov decision processes with uncertain transition rates: Sensitivity and robust control. *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, volume 4, 3799–3804 (IEEE).

Kamrad B, Lele SS, Siddique A, Thomas RJ (2005) Innovation diffusion uncertainty, advertising and pricing policies. *European Journal of Operational Research* 164(3):829–850.

Keskin NB, Zeevi A (2014) Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research* 62(5):1142–1167.

Keskin NB, Zeevi A (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.

Kleywegt AJ, Shapiro A, Homem-de Mello T (2002) The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* 12(2):479–502.

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

Krishnan TV, Bass FM, Kumar V (2000) Impact of a late entrant on the diffusion of a new product/service. *Journal of marketing research* 37(2):269–278.

Kurtz TG (1971) Limit theorems for sequences of jump markov processes approximating ordinary differential processes. *Journal of Applied Probability* 8(2):344–356.

Lenk PJ, Rao AG (1990) New models from old: Forecasting product adoption by hierarchical bayes procedures. *Marketing Science* 9(1):42–53.

Li H (2020) Optimal pricing under diffusion-choice models. *Operations Research* .

Li H, Huh WT (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 AE, Shanthikumar JG (2007) Relative entropy, exponential utility, and robust dynamic pricing. *Operations Research* 55(2):198–214.

Lin KY (2006) Dynamic pricing with real-time demand learning. *European Journal of Operational Research* 174(1):522–538.

Lorentz GG (2013) *Bernstein polynomials* (American Mathematical Soc.).

Mahajan V, Muller E, Bass FM (1995) Diffusion of new products: Empirical generalizations and managerial uses. *Marketing science* 14(3_supplement):G79–G88.

Nilim A, El Ghaoui L (2004) *Robust markov decision processes with uncertain transition matrices*. Ph.D. thesis, University of California, Berkeley.

Nilim A, El Ghaoui L (2005) Robust control of markov decision processes with uncertain transition matrices. *Operations Research* 53(5):780–798.

Niu SC (2002) A stochastic formulation of the Bass model of new-product diffusion. *Mathematical problems in Engineering* 8(3):249–263.

Niu SC (2006) A piecewise-diffusion model of new-product demands. *Operations Research* 54(4):678–695.

Norman MF, et al. (1974) A central limit theorem for markov processes that move by small steps. *The Annals of Probability* 2(6):1065–1074.

Norton JA, Bass FM (1987) A diffusion theory model of adoption and substitution for successive generations of high-technology products. *Management science* 33(9):1069–1086.

Qi A, Ahn HS, Sinha A (2017) Capacity investment with demand learning. *Operations Research* 65(1):145–164.

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

Robinson B, Lakhani C (1975) Dynamic price models for new-product planning. *Management science* 21(10):1113–1122.

Roussas GG (1969) Nonparametric estimation in Markov processes. *Annals of the Institute of Statistical Mathematics* 21(1):73–87.

Savin S, Terwiesch C (2005) Optimal product launch times in a duopoly: Balancing life-cycle revenues with product cost. *Operations Research* 53(1):26–47.

Schmittlein DC, Mahajan V (1982) Maximum likelihood estimation for an innovation diffusion model of new product acceptance. *Marketing science* 1(1):57–78.

Shen W, Duenyas I, Kapuscinski R (2013) Optimal pricing, production, and inventory for new product diffusion under supply constraints. *Manufacturing & Service Operations Management* 16(1):28–45.

Srinivasan V, Mason CH (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. URL `https://stonemaiergames.com/kickstarter-lesson-62-early-bird-pledge-levels`.

Taneja IJ, Kumar P (2004) Relative information of type s, csiszár's f-divergence, and information inequalities. *Information Sciences* 166(1-4):105–125.

USA TODAY JG (2019) Did Apple retail prices get too high in 2018? Consumers say yes. URL `https://www.usatoday.com/story/tech/talkingtech/2018/12/29/did-apple-retail-prices-get-too-high-2018-consumers-say-way-yes/2432445002/`.

Wang Z, Deng S, Ye Y (2014) Close the gaps: A learning-while-doing algorithm for single-product revenue management problems. *Operations Research* 62(2):318–331.

ZDNet AKH (2019) Apple CEO on iPhone sales drop: 'Yes, i do think that price is a factor'. URL https://www.zdnet.com/article/apple-ceo-on-iphone-sales-drop-yes-i-do-think-that-price-is-a-factor/.

Mengzhenyu Zhang is a PhD student in Technology and Operations at the Ross School of Business, University of Michigan. She will join University College London as an assistant professor in 2022. She is interested in studying revenue management and supply chain management problems when interaction among market players impacts market demand.

Hyun-Soo Ahn is a Professor of Technology and Operations and the Ford Motor Co. Director of the Tauber Institute for Global Operations at Ross School of Business at the University of Michigan. His research focuses on developing and analyzing mathematical models related to supply chain management, dynamic pricing and revenue management, workforce agility, and resource allocation.

Joline Uichanco is an Assistant Professor in Technology and Operations at the Ross School of Business, University of Michigan. She does research in analytic models for e-commerce and omnichannel retail operations.

This page is intentionally blank. Proper e-companion title page, with INFORMS branding and exact metadata of the main paper, will be produced by the INFORMS office when the issue is being assembled.

Electronic companion

Table of Contents

EC.1. Algorithms

Algorithm 3 Numerically solve HJB equation (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** $k \le N_{\text{iter}}$ **do** ▷ *Find ν using fixed point iteration*

      $r \leftarrow \inf\left\{r : r \ge -\frac{x(r)}{x'(r)} - V(d+1, t) + \nu_k\right\}$ ▷ *Solve (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 (3.3) for ν*

      $k \leftarrow k + 1$

    $V(d, t) \leftarrow \nu_k$

EC.2. Proofs

EC.2.1. Proof of Proposition 1

Proof. We will use the arguments adapted from Chapter 11 ("Density dependent population processes") in the book Ethier and Kurtz (2005) to prove (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, for any fixed $t \ge 0$,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 $Y := \{Y(t), t \ge 0\}$ 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 $\tilde{Y} := \{Y(t) - t, t \ge 0\}$, which is a "centered" Poisson process with mean zero. Let Z_j^{r,m_0}, $1 \le j \le m_0$, be the jth inter-adoption time in $D^{r,m_0} = \{D_t^{r,m_0}, t \ge 0\}$ 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, for fixed $t \geq 0$, we have

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

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,

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

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

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

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

Therefore, for any $t \geq 0$, by (B.1) and the definition of F_t^r in (2.1), we have

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

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

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

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

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

By Gronwall's inequality, we have

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

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

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

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

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

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

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

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 $1/m_0$.

■

EC.2.2. Proof of Lemma 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 $\zeta_1(t),\zeta_2(t),\dots,\zeta_{m_0}(t)$ 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 $i=1,\dots,m_0$.

Let $\mathcal{X}:=\{\zeta_1(t),\dots,\zeta_{m_0}(t)\}$ 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 $\epsilon>0$,

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

Hence, we have

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

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 (2.3) is the adoption rate function, we have

$$\begin{aligned} f_t^{r,m_0} = \frac{1}{m_0}\mathbb{E}\left[\lambda(D_t^{r,m_0},r_t)\right] &= \mathbb{E}\left[\left(1-\frac{D_t^{r,m_0}}{m_0}\right)\left(p_0+q_0\frac{D_t^{r,m_0}}{m_0}\right)x(r_t)\right] \\ &= \left(p_0\mathbb{E}\left(1-\frac{D_t^{r,m_0}}{m_0}\right)+q_0\mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}-\left(\frac{D_t^{r,m_0}}{m_0}\right)^2\right)\right)x(r_t) \\ &= \left(p_0\left(1-F_t^{r,m_0}\right)+q_0\left(F_t^{r,m_0}-\mathbb{E}\left[\left(\frac{D_t^{r,m_0}}{m_0}\right)^2\right]\right)\right)x(r_t) \\ &= \left(p_0\left(1-F_t^{r,m_0}\right)+q_0\left(F_t^{r,m_0}-\left(F_t^{r,m_0}\right)^2-\mathrm{Var}\left(\frac{D_t^{r,m_0}}{m_0}\right)\right)\right)x(r_t). \end{aligned} \tag{B.10}$$

Dividing both sides of (B.10) by $(1-F_t^{r,m_0})(p_0+q_0F_t^{r,m_0})$, we have

$$\begin{aligned} \frac{f_t^{r,m_0}}{(1-F_t^{r,m_0})(p_0+q_0F_t^{r,m_0})} &= x(r_t)\left[1-\frac{q_0\mathrm{Var}\left(\frac{D_t^{r,m_0}}{m_0}\right)}{(1-F_t^{r,m_0})(p_0+q_0F_t^{r,m_0})}\right] \\ &= x(r_t)\left[1-\frac{q_0}{m_0}\frac{F_t^{r,m_0}+\mathcal{O}(1)\left(\frac{t}{e^t}\right)}{p_0+q_0F_t^{r,m_0}}\right], \end{aligned} \tag{B.11}$$

where the last equality follows from (B.9).

The differential equation (B.11) is similar to the deterministic Bass model (2.1), except with a modified market effort function. Hence, modifying (2.2) results in

$$\mathbb{E}\left(\frac{D_t^{r,m_0}}{m_0}\right) = F_t^{r,m_0} = \frac{1-\exp\left(-(p_0+q_0)\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s\right)}{1+\frac{q_0}{p_0}\exp\left(-(p_0+q_0)\int_0^t\left(1-\frac{q_0}{m_0}\frac{F_s^{r,m_0}+\mathcal{O}(1)\left(\frac{s}{e^s}\right)}{p_0+q_0F_s^{r,m_0}}\right)x(r_s)\mathrm{d}s\right)}. \tag{B.12}$$

Then, from (2.2) and (B.12),

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

where the last inequality follows from Assumption 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 1 (ii), proving that (b) has an upper bound that is $\mathcal{O}(1/m_0)$.

■

EC.2.3. Proof of Theorem 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 (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

$$\frac{\partial V}{\partial t}=\max_{r\in(-\infty,\infty)}\left\{r\lambda(d,r)+\left[V(d+1,t)-V(d,t)\right]\lambda(d,r)\right\}. \tag{B.13}$$

We will show existence of a unique solution $V(\cdot,\cdot)$ to (B.13) at the end of this proof.

We next derive the optimal solution $r^*(d,t)$ to the right-hand side of (B.13). Denote the objective function of the right-hand side of (B.13) by $J(r,d,t)$. Since $\lambda(d,r)=\xi(d)x(r)$, then

$$\frac{\partial J}{\partial r}=\lambda(d,r)x'(r)\left(r+\frac{x(r)}{x'(r)}+V(d+1,t)-V(d,t)\right). \tag{B.14}$$

Note that $x'(r)<0$ (Assumption 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 (3.2). Note that (3.2) has a unique solution. This is because, rearranging (3.2) as

$$0=-r-\frac{x(r)}{x'(r)}-V(d+1,t)+V(d,t), \tag{B.15}$$

the right-hand side is strictly decreasing in r (Assumption 1(iv)), implying that there is a unique root to the equation (B.15).

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

$$\left.\frac{\partial^2 J}{\partial r^2}\right|_{r=r^*}=\frac{\xi(d)}{x'(r)}\left[2x'(r)^2-x(r)x''(r)\right]. \tag{B.16}$$

Since $2x'(r)^2 - x(r)x''(r) \geq 0$ (Assumption 1(iv)), and $x'(r) < 0$ (Assumption 1(iii)), it follows that $\frac{\partial^2 J}{\partial r^2}\Big|_{r=r^*} \leq 0$. Hence $r^*(d,t)$ is the unique maximizer of the right-hand side of the HJB equation (B.13) 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 (B.13) as (3.3).

To complete the proof, we will show that there exists a unique solution V to the HJB equation (B.13). By Theorem VII.T3 (page 208) in Brémaud (1981), a unique solution exists if we can replace $\max\limits_{r\in(-\infty,\infty)}$ in (B.13) 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 (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 (B.13) 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 (B.13) 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 (B.13).

■

EC.2.4. Lemma EC.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 EC.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 (3.3), $\frac{\partial}{\partial t} V(d,t)$ is nonnegative due to the assumption that $x'(r) < 0$ for all r (Assumption 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 (B.13), since $\frac{\partial V(d,t)}{\partial t} = J(r,d,t)\big|_{r=r^*(d,t)}$, we have that

$$\frac{\partial^2 V(d,t)}{\partial d \partial t} = \frac{\partial J(r,d,t)}{\partial r}\frac{\partial r}{\partial d}\Big|_{r=r^*(d,t)} + \frac{\partial J(r,d,t)}{\partial d}\Big|_{r=r^*(d,t)} \tag{B.17}$$

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

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

where the second equality follows from (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

$$\begin{aligned}&\frac{\partial g}{\partial t} + \left[(m_0-d)\left(p_0 + q_0\frac{d}{m_0}\right)x(r^*(d,t))\right]g \\ &\quad= -\frac{x(r^*(d,t))^2}{x'(r^*(d,t))}\left(q_0 - p_0 - 2q_0\frac{d}{m_0}\right) + \frac{\partial V(d+1,t)}{\partial d}(m_0-d)\left(p_0+q_0\frac{d}{m_0}\right)x(r^*(d,t)),\end{aligned} \tag{B.18}$$

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

$$\begin{aligned}&\frac{\partial V(d,t)}{\partial d}\underbrace{e^{(m_0-d)(p_0+\frac{d}{m_0}q_0)(h_1(t)+z_1)}}_{(1)} \\ &= \int_0^t \underbrace{e^{(m_0-d)(p_0+\frac{d}{m_0}q_0)(h_1(s)+z_1)}}_{(2)}\left[\underbrace{-\frac{x(r^*)^2}{x'(r^*)}\left(q_0-p_0-2q_0\frac{d}{m_0}\right)}_{(3)} + \underbrace{(m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)x(r^*)\frac{\partial V(d+1,s)}{\partial d}}_{(4)}\right]\mathrm{d}s,\end{aligned} \tag{B.19}$$

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 $m_0\left(\frac{1}{2}-\frac{p_0}{2q_0}\right) < k = d+1,\dots,m_0-2$. We will then show $\frac{\partial V(d,t)}{\partial d} \le 0$. According to (B.19), $\frac{\partial V(d,t)}{\partial d} \le 0$ because all $(1),(2),(3),(4) \le 0$.

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

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

Hence, taking the partial derivative of (3.3) w.r.t. p_0 and using (B.20), we have

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

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

$$\frac{\partial g}{\partial t}+(m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)x(r^*)g = -\frac{x(r^*)^2}{x'(r^*)}(m_0-d)+(m_0-d)\left(p_0+\frac{d}{m_0}q_0\right)x(r^*)\frac{\partial V(d+1,t)}{\partial p_0}, \tag{B.21}$$

which we solve using the same techniques as (B.18), resulting in

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

As we know $\frac{\partial V(m_0,t)}{\partial p_0} = 0$ for all $t \in [0,T]$, then $\frac{\partial V(m_0-1,t)}{\partial p_0} \ge 0$ 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:

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

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

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 (B.21). Therefore, $V(d,t)$ is partially monotone increasing in p_0, q_0, m_0.

■

EC.2.5. Proof of Corollary 1

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

$$\frac{\partial}{\partial t} V(d,t) = (m_0 - d)\left(p_0 + \frac{d}{m_0} q_0\right) \frac{e^{-2r^*(d,t)}}{e^{-r^*(d,t)}} = \xi(d) e^{V(d+1,t)-V(d,t)-1}. \tag{B.22}$$

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

- From (B.22) 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

$$\frac{\partial V(m_0-1,t)}{\partial t} = \left(\frac{\xi(m_0-1)}{e}\right) \Big/ \left(\frac{\xi(m_0-1)}{e} t + 1\right) \tag{B.23}$$

- From (B.22), and (B.23), 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,

$$\frac{\partial V(m_0-2,t)}{\partial t} = \left(\frac{\xi(m_0-1)\xi(m_0-2)}{e^2} t + \frac{\xi(m_0-2)}{e}\right) \Big/ \left(\frac{\xi(m_0-1)\xi(m_0-2)}{2e^2} t^2 + \frac{\xi(m_0-2)}{e} t + 1\right). \tag{B.24}$$

- From (B.22), 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 (B.23)–(B.24), this reduces to a partial differential equation whose solution is

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

  This then provides us with $\partial V(m_0-3,t)/\partial t$.
- We can continue to solve for $V(0,t)$:

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

■

EC.2.6. Proof of Proposition 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 $\mathbf{z} = (z_1, z_2, z_3)^\top$,

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

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 $\mathbf{z} \neq \mathbf{0}$,

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

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 $\widehat{D}_{t-} \ge 3$ for $\nabla_\beta^2(\sum_{i=0}^{\widehat{D}_{t-}} \ln f_i(\beta)) \prec 0$. Since $\mathcal{L}_t(\widehat{\mathbf{U}}_t |\, \beta) = \sum_{i=0}^{\widehat{D}_{t-}} \ln f_i(\beta)$, we can conclude that $\nabla_\beta^2 \mathcal{L}_t(\widehat{\mathbf{U}}_t |\, \beta) \prec 0$ when $\widehat{D}_{t-} \ge 3$.

■

EC.2.7. Proof of Lemma 3

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

The ML estimators are finite since, from (4.2), if either $\hat{p}_t = +\infty$ or $\hat{q}_t = +\infty$ or $\hat{m}_t = +\infty$, then the likelihood function is 0. Hence, there exist finite $\bar{\delta}_1, \bar{\delta}_2, \bar{\delta}_3$ such that $\hat{p}_t \le p_0(1 + \bar{\delta}_1), \hat{q}_t \le q_0(1 + \bar{\delta}_2)$ and $\hat{m}_t \le m_0(1 + \bar{\delta}_3)$. Note that the ML estimator $\hat{\theta}_t = (\hat{p}_t, \hat{q}_t, \hat{m}_t)$ can be written as

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

where $u = (u_p, u_q, u_m)$, $\theta_0 = (p_0, q_0, m_0)$, and $f_i(\theta)$ is defined in (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 $\bar{\delta}_1 p_0 \geq \delta > 0$. 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,

$$\begin{aligned}
&\mathbb{P}_{\theta_0}\left\{|\hat{p}_t - p_0| > \delta\right\} \\
&\leq \mathbb{P}_{\theta_0}\left\{-\sum_{i=0}^{D_t} \ln \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)} \leq 0\right\} + \mathbb{P}_{\theta_0}\left\{-\sum_{i=0}^{D_t} \ln \frac{f_i(\theta_0 - \delta e_1)}{f_i(\theta_0)} \leq 0\right\} \\
&\leq 2\mathbb{P}_{\theta_0}\left\{-\sum_{i=0}^{D_t} \ln \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)} \leq 0\right\} = 2\mathbb{P}_{\theta_0}\left\{\prod_{i=0}^{D_t} \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)} \geq 1\right\} \\
&= 2\mathbb{P}_{\theta_0}\left\{\sqrt{\prod_{i=0}^{D_t} \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)}} \geq 1\right\} \leq 2\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t} \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)}}\right) \\
&= 2\mathbb{E}_{\theta_0}\left(\mathbb{E}_{\theta_0}\left(\cdots \mathbb{E}_{\theta_0}\left(\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t} \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)}} \mid \mathcal{F}_{t_{D_t-1}}\right) \mid \mathcal{F}_{t_{D_t-2}}\right) \cdots \mid \mathcal{F}_{t_1}\right) \mid \mathcal{F}_0\right).
\end{aligned} \tag{B.25}$$

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 (B.25) starting from the innermost conditional expectation. We have

$$\begin{aligned}
&\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t} \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)}} \mid \mathcal{F}_{t_{D_t-1}}\right) = \sqrt{\prod_{i=0}^{D_t-1} \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)}}\,\mathbb{E}_{\theta_0}\left(\sqrt{\frac{f_{D_t}(\theta_0 + \delta e_1)}{f_{D_t}(\theta_0)}} \mid \mathcal{F}_{t_{D_t-1}}\right) \\
&= \sqrt{\prod_{i=0}^{D_t-1} \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)}}\left(\int_{t_{D_t-1}}^{\infty} \sqrt{\frac{f_{D_t}(\theta_0 + \delta e_1)}{f_{D_t}(\theta_0)}} f_{D_t}(\theta_0)\mathrm{d}t_{D_t}\right) \\
&= \sqrt{\prod_{i=0}^{D_t-1} \frac{f_i(\theta_0 + \delta e_1)}{f_i(\theta_0)}}\left(\int_{t_{D_t-1}}^{\infty} \sqrt{f_{D_t}(\theta_0 + \delta e_1)}\sqrt{f_{D_t}(\theta_0)}\mathrm{d}t_{D_t}\right).
\end{aligned} \tag{B.26}$$

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,

$$\int_{t_{D_t-1}}^{\infty}\sqrt{f_{D_t}(\theta_0+\delta e_1)f_{D_t}(\theta_0)}\mathrm{d}t_{D_t}=1-\frac{1}{2}\int_{t_{D_t-1}}^{\infty}\left(\sqrt{f_{D_t}(\theta_0+\delta e_1)}-\sqrt{f_{D_t}(\theta_0)}\right)^2\mathrm{d}t_{D_t}. \tag{B.27}$$

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) provided the following condition holds. Specifically,

$$\frac{1}{2}\int_{t_{D_t-1}}^{\infty}\left(\sqrt{f_{D_t}(\theta_0+\delta e_1)}-\sqrt{f_{D_t}(\theta_0)}\right)^2\mathrm{d}t_{D_t}\geq\frac{1}{4\sqrt{R}}\mathbb{E}_{\theta_0}\left(\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0+\delta e_1)}\mid\mathcal{F}_{t_{D_t-1}}\right), \tag{B.28}$$

where R is a constant such that $R\geq\max_{\delta\in[0,\bar{\delta}_1p_0],t_{D_t}}\frac{1}{f_{D_t}(\theta_0+\delta e_1)}$. Here, we can choose $R=1/p_0$ since $\max_{\delta\in[0,\bar{\delta}_1p_0],t_{D_t}}\frac{1}{f_{D_t}(\theta_0+\delta e_1)}\leq 1/(m_0p_0)\leq 1/p_0$. Hence, with this choice, R is independent of m_0 and of t. We will next bound the right-hand side of (B.28).

Define $C_I:=(p_0(1+\bar{\delta}_1)+q_0)^2$. Note that

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

where the inequality is because $p_0+\delta\leq p_0(1+\bar{\delta}_1)$.

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

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

Therefore, (B.28) reduces to

$$\frac{1}{4\sqrt{R}C_I}\delta^2 \le \int_{t_{D_t-1}}^{\infty}\left(\sqrt{f_{D_t}(\theta_0+\delta e_1)}-\sqrt{f_{D_t}(\theta_0)}\right)^2 \mathrm{d}t_{D_t}. \tag{B.29}$$

Hence, from (B.27), we have

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

where the first inequality is because $e^{-x}\ge 1-x$ for all x. The second inequality is from (B.29). Hence, from (B.26), we have

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\mid \mathcal{F}_{t_{D_t-1}}\right) \le \sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\theta_0+\delta e_1)}{f_i(\theta_0)}}\cdot\exp\left(-\frac{1}{8\sqrt{R}C_I}\delta^2\right). \tag{B.30}$$

This provides a bound for the innermost conditional expectation in (B.25).

Observe that all the terms in the bound (B.30) 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 (B.30) given $\mathcal{F}_{t_{D_t-2}}$, and using the same logic as the above arguments to bound the right-hand side, we have

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

We can proceed iteratively to evaluate (B.25) as we take conditional expectations given $\mathcal{F}_{t_{D_t}-3}$, $\mathcal{F}_{t_{D_t}-4}$, $\mathcal{F}_0$, resulting in

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

Hence, we have that

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

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

$$\begin{aligned}\mathbb{E}_{\theta_0}\left[(\hat{p}_t-p_0)^2\mid D_t=k\right] &= \int_0^{\infty}\mathbb{P}_{\theta_0}\left\{(\hat{p}_t-p_0)^2>\delta\mid D_t=k\right\}\mathrm{d}\delta = \int_0^{\infty}\mathbb{P}_{\theta_0}\left\{|\hat{p}_t-p_0|^2>\sqrt{\delta}\mid D_t=k\right\}\mathrm{d}\delta\\ &\le \int_0^{\infty}2\exp\left(-\frac{k+1}{16\sqrt{R}C_I}\delta\right)\mathrm{d}\delta = \frac{8\sqrt{R}C_I}{k+1} = \frac{8(p_0(1+\bar{\delta}_1)+q_0)^2}{\sqrt{p_0}}\frac{1}{k+1}.\end{aligned}$$

Thus, we have that $\mathbb{E}_{\theta_0}\left[(\hat{p}_t - p_0)^2/{p_0}^2 \mid D_t = k\right] \le \frac{\alpha_p}{k+1}$ where $\alpha_p := 8\sqrt{R}\left(1+\bar{\delta}_1 + q_0/p_0\right)^2$ is independent of m_0 and of 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

$$\begin{aligned}
\mathbb{E}_{\theta'_0}\left[\frac{\partial^2}{\partial\delta^2}\ln\frac{f_{D_t}(\theta'_0)}{f_{D_t}(\theta'_0+\delta e_2)} \mid \mathcal{F}_{t_{D_t-1}}\right] &= \mathbb{E}_{\theta'_0}\left[\frac{\left(1+\frac{D_t}{m_0}\right)^2}{\left({p_0}' + \left(1+\frac{D_t}{m_0}\right)(q_0+\delta)\right)^2} \mid \mathcal{F}_{t_{D_t-1}}\right] \\
&\ge \mathbb{E}_{\theta'_0}\left[\frac{\left(1+\frac{D_t}{m_0}\right)^2}{\left({p_0}' + \left(1+\frac{D_t}{m_0}\right)(q_0(1+\bar{\delta}_2))\right)^2} \mid \mathcal{F}_{t_{D_t-1}}\right],
\end{aligned}$$

where the inequality is because $q_0 + \delta \le q_0(1+\bar{\delta}_2)$. Defining $C_I := \left(p_0 + q_0(1+\bar{\delta}_2)\right)^2$, we have that

$$\mathbb{E}_{\theta'_0}\left[\frac{\left(1+\frac{D_t}{m_0}\right)^2}{\left(p_0 + \left(1+\frac{D_t}{m_0}\right)(q_0+\bar{\delta}_2)\right)^2} \mid \mathcal{F}_{t_{D_t-1}}\right] \ge \frac{1}{\left(p_0 + q_0(1+\bar{\delta}_2)\right)^2} = \frac{1}{C_I}.$$

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

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

This implies that

$$\begin{aligned}
\mathbb{E}_{\theta_0}\left[(\hat{q}_t - q_0)^2 \mid D_t = k\right] &= \int_0^\infty \mathbb{P}_{\theta_0}\left\{(\hat{q}_t - q_0)^2 > \delta \mid D_t = k\right\} \mathrm{d}\delta = \int_0^\infty \mathbb{P}_{\theta_0}\left\{|\hat{q}_t - q_0|^2 > \sqrt{\delta} \mid D_t = k\right\} \mathrm{d}\delta \\
&\le \int_0^\infty 2\exp\left(-\frac{k+1}{16\sqrt{R}C_I}\delta\right)\mathrm{d}\delta = \frac{8\sqrt{R}C_I}{k+1} = \frac{8\sqrt{R}\left(p_0 + q_0(1+\bar{\delta}_2)\right)^2}{k+1}.
\end{aligned}$$

Thus, we have that $\mathbb{E}_{\theta_0}\left[(\hat{q}_t - q_0)^2/{q_0}^2 \mid D_t = k\right] \le \frac{\alpha_q}{k+1}$ where $\alpha_q := 8\sqrt{R}(p_0/q_0 + 1 + \bar{\delta}_2)^2$ is independent of m_0 and of t.

For $\hat{m}_t$, if we define $C_I := (m_0\bar{\delta}_3)^2/{p_0}^2$, we have

$$\frac{\partial^2}{\partial\delta^2}\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\theta_0+\delta e_3)} \ge \frac{{p_0}^2}{(m_0+\delta-D_t)^2} \ge \frac{{p_0}^2}{(m_0\bar{\delta}_3)^2} = \frac{1}{C_I}.$$

Following the same steps for bounding the mean squared estimation error of $\hat{q}_t$ and $\hat{p}_t$, we know $\mathbb{E}_{\theta_0}\left[(\hat{m}_t - m_0)^2/{m_0}^2 \mid D_t = k\right] \leq \frac{\alpha_m}{k+1}$ where $\alpha_m := 8\sqrt{R}(\bar{\delta}_3)^2/{p_0}^2$ is independent of m_0 and t.

Hence, we prove the lemma with $\alpha_\theta := 24\sqrt{R}\max\left\{(1+\bar{\delta}_1 + q_0/p_0)^2, (p_0/q_0 + 1 + \bar{\delta}_2)^2, \left(\bar{\delta}_3/p_0\right)^2\right\}$, and $R = 1/p_0$. Note that α_θ does not depend on t and m_0. ■

EC.2.8. Lemma EC.2 and proof

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

LEMMA EC.2. *Given any two pricing sample paths $r = (r_t, t \geq 0)$ and $r' = (r'_t, t \geq 0)$ that do not scale up with m_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 $t \geq 0$,*

$$\mathbb{E}_{\theta_0}\left|\frac{D_t^{r,m_0}}{m_0} - \frac{D_t^{r',m_0}}{m_0}\right| = \alpha_1 \frac{t}{e^t}\left|r_t - r'_t\right| + \mathcal{O}\left(\frac{1}{\sqrt{m_0}}\right), \tag{B.31}$$

$$\left|\mathbb{E}_{\theta_0}\left(\frac{D_t^{r,m_0}}{m_0} - \frac{D_t^{r',m_0}}{m_0}\right)\right| = \alpha_2 \frac{t}{e^t}\left|r_t - r'_t\right| \tag{B.32}$$

for some $\alpha_1 > 0, \alpha_2 > 0$ independent of m_0, t, r_t and r'_t.

Observe from Lemma EC.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, $\mathrm{Var}\left(D_t^{r,m_0}/m_0\right)$, decreases in the order of $\mathcal{O}(1/m_0)$ (Proposition 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 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 2. Using the bound on $|F_t^r - F_t^{r'}|$ proves (B.31).

We next prove (B.32). 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 2 in deriving (B.12), 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

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

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 1 (iii). This proves (B.32). ■

EC.2.9. Lemma EC.3 and proof

We next state a result that is also useful for the proof of Proposition 3 and Claim EC.2.

LEMMA EC.3. *Given any two price sample paths $r = (r_t, t \geq 0)$ and $r' = (r_t, t \geq 0)$ that do not scale up with m_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 $t \geq 0$,*

$$\left|\mathbb{E}_{\theta_0}\left(\frac{\xi(D_t^{r,m_0})}{m_0}-\frac{\xi(D_t^{r',m_0})}{m_0}\right)\right| \leq (p_0+q_0)\alpha_1\frac{t}{e^t}|r_t-r_t'|, \tag{B.33}$$

and

$$\left|\mathbb{E}_{\theta_0}\left(\frac{\xi(D_t^{r,m_0})}{\xi(D_t^{r',m_0})}\right)\right| = 1+\alpha_2\frac{t}{e^t}\left|r_t - r_t'\right| \tag{B.34}$$

for some $\alpha_1 > 0, \alpha_2 > 0$ *independent of* m_0, t, r_t *and* r_t'.

Proof. Using the definition of $\xi(\cdot)$ in (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

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

Then, (B.33) follows from Lemma EC.2.

To prove (B.34), 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

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

where the second inequality follows from the fact that when $x > y \ge 0$ and $\max\{c_1,c_2\}\ge 0$, $c_1x - c_2y \le \max\{c_1,c_2\}(x-y)$, and the last equality is from Lemma EC.2 and $\exp\left(\frac{t}{e^t}\mathcal{O}\left(\left|r_t-r_t'\right|\right)\right) = 1+\frac{t}{e^t}\mathcal{O}\left(\left|r_t-r_t'\right|\right)$. ■

EC.2.10. Proof of Proposition 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 $d \le m_0 - 1$, we can write the expected revenue-to-go as

$$\begin{aligned} v(d, n\delta t) &= r^\pi(d, n\delta t) \cdot \xi(d) x(r^\pi(d, n\delta t))\delta t \\ &\quad + [v(d+1, (n-1)\delta t) - v(d, (n-1)\delta t)] \cdot \xi(d) x(r^\pi(d, n\delta t))\delta t + v(d, (n-1)\delta t) + o(\delta t), \end{aligned} \tag{B.35}$$

where the adoption probability $\xi(d)x(r)\delta t$ is under a Markovian Bass demand model with parameter vector θ_0 and $o(\delta t)$ is a term such that $\lim_{\delta t \to 0} o(\delta t)/\delta t = 0$. 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 oracle policy π^* which knows the true value θ_0. For any $d \le m_0 - 1$, $V(d, n\delta t)$ can be expressed as

$$\begin{aligned} V(d, n\delta t) &= r^*(d, n\delta t) \cdot \xi(d) x(r^*(d, n\delta t))\delta t \\ &\quad + [V(d+1, (n-1)\delta t) - V(d, (n-1)\delta t)] \cdot \xi(d) x(r^*(d, n\delta t))\delta t + V(d, (n-1)\delta t) + o(\delta t), \end{aligned} \tag{B.36}$$

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

Let $D^\pi = (D^\pi_t, t \ge 0)$ and $D^* = (D^*_t, t \ge 0)$ be the cumulative demand process under π and π^*, respectively. Let $(r^\pi_t, t \ge 0)$ and $(r^*_t, t \ge 0)$ 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 $n = 0, 1, \ldots, N-1$,

$$\Psi_n = \mathcal{O}\left(\mathbb{E}_{\theta_0}\left[\sum_{s=n}^{N-1} \frac{D^\pi_{s\delta t} + 1}{s\delta t + t_0}(r^\pi_{s\delta t} - r^*_{s\delta t})^2 \delta t \mid \mathcal{F}_{n\delta t}\right] + (N-n) \cdot o(\delta t)\right). \tag{B.37}$$

Here, $\mathcal{O}$ describes the limiting behavior as m_0 grows, and the terms inside $\mathcal{O}$ are potentially affected by m_0. Proposition 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 (B.37), 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

$$\begin{aligned} V(d^*_{\omega,T-\delta t}, \delta t) - v(d^\pi_{\omega,T-\delta t}, \delta t) &= \rho^*_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^*_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^\pi_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t + o(\delta t) \\ &= \rho^*_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^*_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t \\ &\quad + \underbrace{\rho^\pi_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^\pi_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t}_{\text{(A')}} + o(\delta t). \end{aligned} \tag{B.38}$$

Note that

$$\begin{aligned} \text{(A')} &= \frac{\rho^\pi_{\omega,T-\delta t}x(\rho^\pi_{\omega,T-\delta t})}{\rho^*_{\omega,T-\delta t}x(\rho^*_{\omega,T-\delta t})} V(d^*_{\omega,T-\delta t}, \delta t) - v(d^\pi_{\omega,T-\delta t}, \delta t) \\ &\leq \frac{\rho^\pi_{\omega,T-\delta t}x(\rho^\pi_{\omega,T-\delta t})}{\rho^*_{\omega,T-\delta t}x(\rho^*_{\omega,T-\delta t})} V(d^*_{\omega,T-\delta t}, \delta t) - \frac{\rho^\pi_{\omega,T-\delta t}x(\rho^\pi_{\omega,T-\delta t})}{\rho^*_{\omega,T-\delta t}x(\rho^*_{\omega,T-\delta t})} v(d^\pi_{\omega,T-\delta t}, \delta t) \\ &= \frac{\rho^\pi_{\omega,T-\delta t}x(\rho^\pi_{\omega,T-\delta t})}{\rho^*_{\omega,T-\delta t}x(\rho^*_{\omega,T-\delta t})} \left(V(d^*_{\omega,T-\delta t}, \delta t) - v(d^\pi_{\omega,T-\delta t}, \delta t)\right) \end{aligned}$$

where the inequality comes from $0 < \frac{\rho^\pi_{\omega,T-\delta t}x(\rho^\pi_{\omega,T-\delta t})}{\rho^*_{\omega,T-\delta t}x(\rho^*_{\omega,T-\delta t})} \leq 1$. Therefore, we can bound (B.38) as follows:

$$\begin{aligned} \text{(B.38)} \leq & \rho^*_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^*_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t \\ & + \frac{\rho^\pi_{\omega,T-\delta t}x(\rho^\pi_{\omega,T-\delta t})}{\rho^*_{\omega,T-\delta t}x(\rho^*_{\omega,T-\delta t})} \left(V(d^*_{\omega,T-\delta t}, \delta t) - v(d^\pi_{\omega,T-\delta t}, \delta t)\right) + o(\delta t), \end{aligned}$$

which implies

$$V(d^*_{\omega,T-\delta t}, \delta t) - v(d^\pi_{\omega,T-\delta t}, \delta t) \leq \frac{1}{1 - \frac{\rho^\pi_{\omega,T-\delta t}x(\rho^\pi_{\omega,T-\delta t})}{\rho^*_{\omega,T-\delta t}x(\rho^*_{\omega,T-\delta t})}} \underbrace{\left|\rho^*_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^*_{\omega,T-\delta t})\delta t - \rho^\pi_{\omega,T-\delta t}\xi(d^*_{\omega,T-\delta t})x(\rho^\pi_{\omega,T-\delta t})\delta t\right|}_{\text{(A)}} + o(\delta t).$$

We examine (A) as follows.

- **Bounding (A):** Recall that we have proved in Theorem 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

$$\left.\frac{\partial[r\xi(d^*_{\omega,T-\delta t})x(r)\delta t]}{\partial r}\right|_{r=\rho^*_{\omega,T-\delta t}} = 0. \tag{B.39}$$

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

$$
\begin{aligned}
(\mathrm{A}) &= \left|\int_{\rho^{\pi}_{\omega,T-\delta t}}^{\rho^{*}_{\omega,T-\delta t}} \frac{\partial[r\xi(d^{*}_{\omega,T-\delta t})x(r)\delta t]}{\partial r}\mathrm{d}r\right| = \left|\int_{\rho^{\pi}_{\omega,T-\delta t}}^{\rho^{*}_{\omega,T-\delta t}} \left(\frac{\partial[r\xi(d^{*}_{\omega,T-\delta t})x(r)\delta t]}{\partial r} - 0\right)\mathrm{d}r\right| \\
&= \left|\int_{\rho^{\pi}_{\omega,T-\delta t}}^{\rho^{*}_{\omega,T-\delta t}} \left(\frac{\partial[r\xi(d^{*}_{\omega,T-\delta t})x(r)\delta t]}{\partial r} - \frac{\partial[r\xi(d^{*}_{\omega,T-\delta t})x(r)\delta t]}{\partial r}\Big|_{r=\rho^{*}_{\omega,T-\delta t}}\right)\mathrm{d}r\right| \qquad (B.40) \\
&= \left|\int_{\rho^{\pi}_{\omega,T-\delta t}}^{\rho^{*}_{\omega,T-\delta t}} \int_{\rho^{*}_{\omega,T-\delta t}}^{r} \frac{\partial^2[z\xi(d^{*}_{\omega,T-\delta t})x(z)\delta t]}{\partial z^2}\mathrm{d}z\mathrm{d}r\right| \\
&\le \xi(d^{*}_{\omega,T-\delta t})\delta t \cdot \sup_{z\in(-\infty,\infty)}\left|\frac{\partial^2}{\partial z^2}(zx(z))\right| \cdot \left|\int_{\rho^{\pi}_{\omega,T-\delta t}}^{\rho^{*}_{\omega,T-\delta t}} \int_{\rho^{*}_{\omega,T-\delta t}}^{r} \mathrm{d}z\mathrm{d}r\right| \\
&\le \frac{1}{2}C_{xx}\xi(d^{*}_{\omega,T-\delta t})(\rho^{\pi}_{\omega,T-\delta t} - \rho^{*}_{\omega,T-\delta t})^2\delta t. \qquad (B.41)
\end{aligned}
$$

Here, (B.40) comes from (B.39), while (B.41) comes from Assumption 1(v).

We replace $\xi(d^{*}_{\omega,T-\delta t})$ in (B.41) 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 EC.3, (B.41) is bounded above by

$$
\begin{aligned}
&\frac{1}{2}C_{xx}\left(1+\frac{T-\delta t}{e^{T-\delta t}}\mathcal{O}\left(\left|\rho^{\pi}_{\omega,T-\delta t}-\rho^{*}_{\omega,T-\delta t}\right|\right)\right)\xi(d^{\pi}_{\omega,T-\delta t})(\rho^{\pi}_{\omega,T-\delta t}-\rho^{*}_{\omega,T-\delta t})^2\delta t \\
&= \frac{1}{2}C_{xx}\left(1+\frac{1}{(T-\delta t)^3}o\left(\left|\rho^{\pi}_{\omega,T-\delta t}-\rho^{*}_{\omega,T-\delta t}\right|\right)\right)\xi(d^{\pi}_{\omega,T-\delta t})(\rho^{\pi}_{\omega,T-\delta t}-\rho^{*}_{\omega,T-\delta t})^2\delta t. \qquad (B.42)
\end{aligned}
$$

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

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

$$
\mathbb{E}(\xi(D^{\pi}_t)\mid\mathcal{F}_t) \le \alpha_1\left(\frac{\mathbb{E}\left[\int_0^{t+t_0}\xi(D^{\pi}_s)x(r^{\pi}_s)\mathrm{d}s\mid\mathcal{F}_t\right]}{t+t_0}\right) = \alpha_2\left(\frac{D^{\pi}_t+1}{t+t_0}\right) \qquad (B.43)
$$

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 $0<h\le 1/m_0$, we have $p_0\le\xi(j)\le m_0\frac{(p_0+q_0)^2}{4q_0}$, which implies $\xi(D^{\pi}_t)\le\Theta(m_0)$ 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 $0\le s\le t$,

$$
\xi(D^{\pi}_s)\ge\min\{\xi(0),\xi(D^{\pi}_t)\}=\min\{m_0p_0,\xi(D^{\pi}_t)\}.
$$

Therefore,

$$
\int_0^t\xi(D^{\pi}_s)x(r^{\pi}_s)\mathrm{d}s\ge\min\{m_0p_0,\xi(D^{\pi}_t)\}\int_0^t x(r^{\pi}_s)\mathrm{d}s\ge\alpha_3(\xi(D^{\pi}_t))\int_0^t x(r^{\pi}_s)\mathrm{d}s\ge\alpha_4(\xi(D^{\pi}_t)t)
$$

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

$$\mathbb{E}(\xi(D_t^\pi) \mid \mathcal{F}_t) \le \alpha_1 \left(\frac{\mathbb{E}\left[\int_0^t \xi(D_s^\pi) x(r_s^\pi) \mathrm{d}s \mid \mathcal{F}_t\right]}{t} \right)$$

with $\alpha_1 = 1/\alpha_4$, which gives us (B.43). Note that we use $D_t^\pi + 1$ 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. Taking the conditional expectation of (B.42) given $\mathcal{F}_{T-\delta t}$ and given the fact that $\rho^\pi_{\omega,T-\delta t}, \rho^*_{\omega,T-\delta t}$ are prices that do not scale up with m_0, we have

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

The last relation is due to Claim EC.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 (B.37) holds for $n+1$. Specifically,

$$\begin{aligned}\Psi_{n+1} &:= \mathbb{E}\left(\left|V(D^*_{(n+1)\delta t}, T-(n+1)\delta t) - v(D^\pi_{(n+1)\delta t}, T-(n+1)\delta t)\right| \mid \mathcal{F}_{(n+1)\delta t}\right) \\ &= \mathcal{O}\left(\mathbb{E}_{\theta_0}\left[\sum_{s=n+1}^{N-1} \frac{D^\pi_{s\delta t}+1}{s\delta t + t_0}(\rho^\pi_{\omega,s\delta t} - \rho^*_{\omega,s\delta t})^2 \delta t \mid \mathcal{F}_{(n+1)\delta t}\right] + (N-n-1)\cdot o(\delta t)\right)\end{aligned} \tag{B.44}$$

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

$$\begin{aligned}&V\left(d^*_{\omega,n\delta t}, T-n\delta t\right) - v\left(d^\pi_{\omega,n\delta t}, T-n\delta t\right) \\ &\quad = \rho^*_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t - \rho^\pi_{\omega,n\delta t}\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\ &\qquad + \left[V(d^*_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^*_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t \\ &\qquad - \left[v(d^\pi_{\omega,n\delta t}+1, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^\pi_{\omega,n\delta t})x(\rho^\pi_{\omega,n\delta t})\delta t \\ &\qquad + V(d^*_{\omega,n\delta t}, T-(n+1)\delta t) - v(d^\pi_{\omega,n\delta t}, T-(n+1)\delta t) + o(\delta t) \\ &\quad \le (\mathrm{A}) + (\mathrm{B}) + (\mathrm{C}) + o(\delta t)\end{aligned} \tag{B.45}$$

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

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

and

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

We let

$$V' := \rho^{\pi}_{\omega,n\delta t}\xi(d^{*}_{\omega,n\delta t})x(\rho^{\pi}_{\omega,n\delta t})\delta t + \left[V(d^{*}_{\omega,n\delta t}+1, T-(n+1)\delta t) - V(d^{*}_{\omega,n\delta t}, T-(n+1)\delta t)\right]\xi(d^{*}_{\omega,n\delta t})x(\rho^{\pi}_{\omega,n\delta t})\delta t.$$

Then, we have

$$\begin{aligned}
(\mathrm{B}) = &\frac{V'}{V\left(d^{*}_{\omega,n\delta t}, T-n\delta t\right)} V\left(d^{*}_{\omega,n\delta t}, T-n\delta t\right) - v\left(d^{\pi}_{\omega,n\delta t}, T-n\delta t\right) \\
\leq &\frac{V'}{V\left(d^{*}_{\omega,n\delta t}, T-n\delta t\right)}\left(V\left(d^{*}_{\omega,n\delta t}, T-n\delta t\right) - v\left(d^{\pi}_{\omega,n\delta t}, T-n\delta t\right)\right)
\end{aligned} \tag{B.46}$$

where the inequality comes from the fact that $0 < \frac{V'}{V\left(d^{*}_{\omega,n\delta t}, T-n\delta t\right)} \leq 1$. Taking (B.46) into (B.45), we know

$$V\left(d^{*}_{\omega,n\delta t}, T-n\delta t\right) - v\left(d^{\pi}_{\omega,n\delta t}, T-n\delta t\right) \leq \frac{1}{1-\frac{V'}{V\left(d^{*}_{\omega,n\delta t}, T-n\delta t\right)}}\left(|(\mathrm{A})| + |(\mathrm{C})|\right) + o(\delta t). \tag{B.47}$$

We will bound $|(\mathrm{A})|$, and $|(\mathrm{C})|$ separately.

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

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

The only difference between (A1) and (A2) is $\rho^{*}_{\omega,n\delta t}$ and $\rho^{\pi}_{\omega,n\delta t}$. Recall from Theorem 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 (B.41), 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 1) and from Assumption 1(v). Note that from Assumption 1(iv), $\frac{x(r)}{x'(r)} + r$ is finite since r is finite.

Hence, $|(\mathrm{A})|$ is upper bounded by

$$\mathcal{O}\left(\xi(d^*_{\omega,n\delta t})(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t})^2\delta t\right) = \mathcal{O}\left(\xi(d^\pi_{\omega,n\delta t})(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t})^2\delta t\right). \tag{B.48}$$

Here the equality comes from the same argument as when we bounded (B.41) with (B.42).

From (B.47), and from the bound (B.48), the following constraint must hold almost surely:

$$\begin{aligned}
&|V(D^*_{n\delta t}, T-(n+1)\delta t) - v(D^\pi_{n\delta t}, T-(n+1)\delta t)| \\
&\le \mathcal{O}\left(\xi(D^\pi_{n\delta t})(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t})^2\delta t\right) \\
&\quad + \mathcal{O}\left(|V(D^*_{n\delta t}, T-(n+1)\delta t) - v(D^\pi_{n\delta t}, T-(n+1)\delta t)|\right) + o(\delta t).
\end{aligned}$$

Therefore, taking the conditional expectation of the above bound given $\mathcal{F}_{n\delta t}$,

$$\begin{aligned}
\Psi_n &\le \mathcal{O}\left(\mathbb{E}\left[\xi(D^\pi_{n\delta t}) \mid \mathcal{F}_{n\delta t}\right](r^*_{n\delta t} - r^\pi_{n\delta t})^2\delta t\right) + \mathbb{E}\left[\Psi_{n+1} \mid \mathcal{F}_{n\delta t}\right] + o(\delta t) \\
&\le \mathcal{O}\left(\frac{D^\pi_{n\delta t}+1}{n\delta t + t_0}\left(\rho^*_{\omega,n\delta t} - \rho^\pi_{\omega,n\delta t}\right)^2\delta t\right) + \mathbb{E}\left[\Psi_{n+1} \mid \mathcal{F}_{n\delta t}\right] + o(\delta t) \\
&= \mathcal{O}\left(\mathbb{E}\left[\sum_{s=n}^{N-1}\frac{D^\pi_{s\delta t}+1}{s\delta t + t_0}\left(\rho^\pi_{\omega,s\delta t} - \rho^*_{\omega,s\delta t}\right)^2\delta t \mid \mathcal{F}_{n\delta t}\right] + (N-n)\cdot o(\delta t)\right)
\end{aligned} \tag{B.49}$$

Here, the second inequality is due to Claim EC.1. The last step is due to the inductive assumption. This finishes the induction proof, thus proving (B.37).

Note that (B.37) is true for any $\delta t > 0$ and for any $n \in \{0, 1, \ldots, N-1\}$. Hence, setting $n = 0$ and taking the limit on both sides of (B.37) as δt goes to zero, we have:

$$\begin{aligned}
\lim_{\delta t\to 0}\Psi_0 &= \lim_{\delta t\to 0}\mathcal{O}\left(\mathbb{E}\left[\sum_{s=0}^{N-1}\frac{D^\pi_{s\delta t}+1}{s\delta t + t_0}\left(\rho^\pi_{\omega,s\delta t} - \rho^*_{\omega,s\delta t}\right)^2\delta t \mid \mathcal{F}_0\right] + N\delta t\cdot\frac{o(\delta t)}{\delta t}\right) \\
&= \mathcal{O}\left(\mathbb{E}\left[\int_{t=0}^{T}\frac{D^\pi_t+1}{t+t_0}\left(\rho^\pi_{\omega,t} - \rho^*_{\omega,t}\right)^2\mathrm{d}t \mid \mathcal{F}_0\right]\right),
\end{aligned}$$

where the last relation follows because $N\delta t = T$ and since $\lim_{\delta t\to 0} o(\delta t)/\delta t = 0$. ■

EC.2.11. Claim EC.2 and proof

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

Claim EC.2. *Under Assumption 2, the upper bound in Proposition 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 of Claim EC.2 Using similar logic as the proof of Proposition 3, we use induction to prove the more general result on a discretized time horizon:

$$\Psi_n = \Omega\left(\mathbb{E}\left[\sum_{s=n}^{N-1} \xi(D^*_{s\delta t})(r^\pi_{s\delta t} - r^*_{s\delta t})^2 \delta t \mid \mathcal{F}_{n\delta t}\right]\right). \tag{B.50}$$

for all $n = 0, 1, \ldots, N-1$. Note that $\Psi_0 = R^* - R(\pi)$. We need to revise the proof of Proposition 3 in several steps to show (B.50). Following the logic of the proof of Proposition 3, we first consider the base step $n = N-1$ with δt time remaining. Recall from (B.38) that

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

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

$$\begin{aligned}
\text{(A)} &= \left|\int_{\rho^\pi_{\omega,T-\delta t}}^{\rho^*_{\omega,T-\delta t}} \int_{\rho^*_{\omega,T-\delta t}}^{r} \frac{\partial^2[z\xi(d^*_{\omega,T-\delta t})x(z)\delta t]}{\partial z^2} \mathrm{d}z\mathrm{d}r\right| \\
&\ge \frac{1}{2}\underline{C}\xi(d^*_{\omega,T-\delta t})(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t})^2\delta t = \Theta\left(\xi(d^*_{\omega,T-\delta t})(\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t})^2\delta t\right).
\end{aligned} \tag{B.51}$$

The equality follows from the arguments in (B.41), and the inequality follows from Assumption 2. Also, from Lemma EC.3, we know that

$$\text{(B)} \le C_x \left|\left(1 + \frac{1}{(T-\delta t)^3} o\left(\left|\rho^\pi_{\omega,T-\delta t} - \rho^*_{\omega,T-\delta t}\right|\right)\right) \frac{T-\delta t}{e^{T-\delta t}} \mathcal{O}\left(\xi(d^\pi_{\omega,T-\delta t})\right)\delta t\right|,$$

which diminishes fast.

Therefore, combining the arguments above, we have

$$\begin{aligned}
\Psi_{N-1} &= \mathbb{E}\left[\left|V(D^*_{T-\delta t}, \delta t) - v(D^\pi_{T-\delta t}, \delta t)\right|\right] \ge \mathbb{E}\left[\text{(A)}\right] - \mathbb{E}\left[\text{(B)}\right] \\
&\ge \Omega\left(\mathbb{E}\left[\xi(D^*_{T-\delta t})(r^\pi_{T-\delta t} - r^*_{T-\delta t})^2\delta t\right]\right) - \frac{T-\delta t}{e^{T-\delta t}} \cdot \mathcal{O}(\mathbb{E}\left[\xi(D^\pi_{T-\delta t})\right])\delta t \\
&= \Omega\left(\mathbb{E}\left[\xi(D^*_{T-\delta t})(r^\pi_{T-\delta t} - r^*_{T-\delta t})^2\delta t\right]\right).
\end{aligned}$$

This finishes the base step.

For the inductive step, we assume that the result holds for $n+1$. Specifically,

$$\Psi_{n+1} = \Omega\left(\mathbb{E}\left[\sum_{s=n+1}^{N-1} \xi(D^*_{s\delta t})(r^\pi_{s\delta t} - r^*_{s\delta t})^2 \delta t \mid \mathcal{F}_{(n+1)\delta t}\right]\right). \tag{B.52}$$

We need to show that this implies the result holding for n.

We revise the proof of Proposition 3 as follows. First, recall from (B.47), we have

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

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 $|x+y|\ge|x|-|y|$, we know

$$|V(d^*_{\omega,n\delta t},T-n\delta t)-v(d^\pi_{\omega,n\delta t},T-n\delta t)|=|(\text{C"})+(\text{A"})+(\text{B"})|\ge|(\text{C"})+(\text{A"})|-|(\text{B"})|. \tag{B.53}$$

Note that (A") ≥ 0 and (C") ≥ 0 because (A") is the difference between the expected revenue during δt under optimal price $\rho^*_{\omega,n\delta t}$

$$\rho^*_{\omega,n\delta t}\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t+\left[V(d^*_{\omega,n\delta t}+1,T-(n+1)\delta t)-V(d^*_{\omega,n\delta t},T-(n+1)\delta t)\right]\xi(d^*_{\omega,n\delta t})x(\rho^*_{\omega,n\delta t})\delta t$$

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 sub-optimal price path. Hence, we know the right hand side of (B.53) equals to (C") + (A") − |(B")|.

We know the following holds from the definition of π^*:

$$\begin{aligned}
|(\text{A"})|&=\xi(d^*_{\omega,n\delta t})\delta t\cdot\left|\int_{\rho^\pi_{\omega,n\delta t}}^{\rho^*_{\omega,n\delta t}}\int_{\rho^*_{\omega,n\delta t}}^{r}\left(\frac{\partial^2 zx(z)}{\partial z^2}+[V(d^*_{\omega,n\delta t}+1,T-(n+1)\delta t)-V(d^*_{\omega,n\delta t},T-(n+1)\delta t)]x''(z)\right)\mathrm{d}z\mathrm{d}r\right| \\
&\ge\xi(d^*_{\omega,n\delta t})\delta t\cdot\inf_z\left|\frac{\partial^2}{\partial z^2}\left(zx(z)\right)-\left(\frac{x(\bar r)}{x'(\bar r)}+\bar r\right)x''(z)\right|\cdot\int_{\rho^\pi_{\omega,n\delta t}}^{\rho^*_{\omega,n\delta t}}\int_{\rho^*_{\omega,n\delta t}}^{r}\mathrm{d}z\mathrm{d}r,
\end{aligned}$$

where $\bar{r}$ is the price that optimizes the expected revenue-to-go given state $(d^*_{\omega,n\delta t}, T-(n+1)\delta t)$. Hence, according to Assumption 1 and Assumption 2, we know

$$|(\text{A"})| = \Theta\left(\xi(d^*_{\omega,n\delta t})(\rho^*_{\omega,n\delta t}-\rho^\pi_{\omega,n\delta t})^2\delta t\right) \tag{B.54}$$

Therefore, together with (B.52) and (B.54), we know from taking the conditional expectation of (B.53) given $\mathcal{F}_{n\delta t}$ that:

$$\begin{aligned}\Psi_n &= \Omega\left(\mathbb{E}\left[|V(D^*_{n\delta t}, T-(n+1)\delta t) - v(D^\pi_{n\delta t}, T-(n+1)\delta t)| + \xi(D^*_{n\delta t})\left(r^*_{n\delta t}-r^\pi_{n\delta t}\right)^2\delta t \mid \mathcal{F}_{n\delta t}\right]\right)\\ &= \Omega\left(\mathbb{E}\left[\sum_{s=n}^{N-1}\xi(D^*_{s\delta t})\left(r^\pi_{s\delta t}-r^*_{s\delta t}\right)\delta t \mid \mathcal{F}_{n\delta t}\right]\right),\end{aligned}$$

where the last equation follows from the inductive hypothesis.

end proof of Claim EC.2 ■

EC.2.12. Proof of Lemma 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 $\bar{\theta}_t = (\bar{p}_t, \bar{q}_t, \bar{m}_t)$ 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:

$$\left(r^*_t(\bar{\theta}_t, D^\pi_t) - r^*_t(\theta_0, D^*_t)\right)^2 = \Theta\left(\underbrace{\left[r^*_t(\bar{\theta}_t, D^\pi_t) - r^*_t(\theta_0, D^\pi_t)\right]^2}_{\text{(A)}}\right) + \Theta\left(\underbrace{\left[r^*_t(\theta_0, D^\pi_t) - r^*_t(\theta_0, D^*_t)\right]^2}_{\text{(B)}}\right).$$

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 $\bar{p} := \max\{p_0, \bar{p}_t\}$ and $\underline{p} := \min\{p_0, \bar{p}_t\}$. Since $p_0, \bar{p}_t$ 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 $|x-y|^\top (\inf_z \nabla h(z)) \le h(x) - h(y) \le |x-y|^\top (\sup_z \nabla h(z))$, we have that

$$\text{(A)} \le \Theta\left[\left(\sup_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r^*_t}{\partial p}\right|\right)^2 (p_0-\bar{p}_t)^2 + \left(\sup_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r^*_t}{\partial q}\right|\right)^2 (q_0-\bar{q}_t)^2 + \left(\sup_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r^*_t}{\partial m}\right|\right)^2 (m_0-\bar{m}_t)^2\right],$$

and

$$\text{(A)} \ge \Theta\left[\left(\inf_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r^*_t}{\partial p}\right|\right)^2 (p_0-\bar{p}_t)^2 + \left(\inf_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r^*_t}{\partial q}\right|\right)^2 (q_0-\bar{q}_t)^2 + \left(\inf_{(p,q,m)\in\mathcal{P}}\left|\frac{\partial r^*_t}{\partial m}\right|\right)^2 (m_0-\bar{m}_t)^2\right],$$

Here we treat m as a continuous variable.

We first analyze $|\partial r_t^*/\partial p|$. Using the equation (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,

$$\left|\frac{\partial r_t^*(\theta,d)}{\partial p}\right| = \left|\frac{\partial}{\partial p}[V(d,T-t)-V(d+1,T-t)]\Big/\left(\frac{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}{x'(r_t^*)^2}\right)\right| \tag{B.55}$$

Note that if we rearrange (B.21), where $g(d,t)=\frac{\partial V(d,t)}{\partial p}$, we have

$$\underbrace{\frac{\partial}{\partial p}[V(d,T-t)-V(d+1,T-t)]}_{\text{(A1)}} = \underbrace{\frac{\partial^2 V(d,T-t)}{\partial p\partial t}\cdot\frac{1}{(m-d)\left(p+\frac{d}{m}q\right)x(r_t^*)}}_{\text{(A2)}} - \frac{x(r_t^*)}{x'(r_t^*)}\frac{1}{p+\frac{d}{m}q}. \tag{B.56}$$

We now examine (A2) on the right-hand side of (B.56). From the HJB equation (B.13), 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

$$\frac{\partial^2 V(d,T-t)}{\partial p\partial t} = -\frac{\partial J(r,d,T-t)}{\partial r}\Big|_{r=r_t^*}\frac{\partial r_t^*}{\partial p} - \frac{\partial J(r,d,T-t)}{\partial p}\Big|_{r=r_t^*} = 0 - \frac{\partial J(r,d,T-t)}{\partial p}\Big|_{r=r_t^*} \le 0.$$

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

$$\frac{\partial^2 V(d,T-t)}{\partial p\partial t} \ge -(m-d)r_t^* x(r_t^*).$$

Hence, from these lower and upper bounds that we derived for $\frac{\partial^2 V(d,T-t)}{\partial p\partial t}$, (B.56) implies that

$$-\frac{r_t^*}{p+\frac{d}{m}q} - \frac{x(r_t^*)}{x'(r_t^*)}\frac{1}{p+\frac{d}{m}q} \le \text{(A1)} \le -\frac{x(r_t^*)}{x'(r_t^*)}\frac{1}{p+\frac{d}{m}q}. \tag{B.57}$$

Therefore, we substitute (B.57) into (B.55) to get

$$\left|\frac{\partial r_t^*}{\partial p}\right| \le \left|\frac{x'(r_t^*)^2}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right|\cdot\left(\left|\frac{x(r_t^*)}{x'(r_t^*)}\frac{1}{p+\frac{d}{m}q}\right| + \left|\frac{r_t^*}{p+\frac{d}{m}q}\right|\right) \le \Theta\left(\frac{M\bar{x}^u}{C_d(\underline{p}+\underline{q})}\right),$$

where the last inequality follows from Assumption 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 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

$$\left|\frac{\partial r_t^*}{\partial q}\right| \le \left|\frac{x'(r_t^*)^2}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right|\cdot\left(\left|\frac{x(r_t^*)}{x'(r_t^*)}\frac{\frac{d}{m}}{p+\frac{d}{m}q}\right| + \left|\frac{r_t^*}{p+\frac{d}{m}q}\right|\right) \le \Theta\left(\frac{M\bar{x}^u}{C_d(\underline{p}+\underline{q})}\right),$$

and

$$\left|\frac{\partial r_t^*}{\partial m}\right| \le \left|\frac{x'(r_t^*)^2}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right|\cdot\left(\left|\frac{x(r_t^*)}{x'(r_t^*)}\frac{p+\frac{d^2}{m^2}q}{(m-d)(p+\frac{d}{m}q)}\right| + \left|\frac{r_t^*}{p+\frac{d}{m}q}\right|\right) \le \Theta\left(\frac{M\bar{x}^u}{C_d}\frac{p+q}{p\underline{m}}\right) \le \Theta\left(\frac{M\bar{x}^u}{C_d}\frac{p+q}{pm_0}\right).$$

Hence, it follows that

$$(\mathrm{A}) = \Theta\left(\left(\frac{\bar{p}_t - p_0}{p_0}\right)^2 + \left(\frac{\bar{q}_t - q_0}{q_0}\right)^2 + \left(\frac{\bar{m}_t - m_0}{m_0}\right)^2\right).$$

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

$$\begin{aligned} &\mathbb{E}_{\theta_0}\left[|r_t^*(\theta_0, D_t^*) - r_t^*(\theta_0, D_t^\pi)| \mid \mathcal{F}_t\right] \\ &\le \mathbb{E}_{\theta_0}\left[\sup_{d\in[D_t^*\wedge D_t^\pi,\, D_t^*\vee D_t^\pi]}\left|\frac{\partial r_t^*(\theta_0, d)}{\partial f_d}\right| \cdot \sup_{d\in[D_t^*\wedge D_t^\pi,\, D_t^*\vee D_t^\pi]}\left|\frac{\partial f_d}{\partial (d/m_0)}\right| \cdot \left|\frac{D_t^\pi}{m_0} - \frac{D_t^*}{m_0}\right| \mid \mathcal{F}_t\right]. \end{aligned} \tag{B.58}$$

Note that $\sup_{d\in[0,m_0]} |\partial f_d / \partial (d/m_0)| = (p_0 + q_0)$. Hence, to bound (B.58), we need to evaluate the bound of $|\partial r_t^* / \partial f_d|$.

From (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 (B.12), 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

$$\mathbb{E}_{\theta_0}\left[f_{D_t^*} \mid \mathcal{F}_t\right] = \mathbb{E}_{\theta_0}\left[\left(1 - \frac{D_t^*}{m_0}\right)\left(p_0 + q_0 \frac{D_t^*}{m_0}\right) \mid \mathcal{F}_t\right] > \gamma^\delta. \tag{B.59}$$

For any (d, t), we differentiate (3.3) by f_d for both sides, which yields

$$\frac{\partial^2 V(d, T-t)}{\partial f_d \partial t} \Big/ m_0 + \frac{x(r_t^*)^2}{x'(r_t^*)} + f_d \frac{2x(r_t^*)x'(r_t^*)^2 - x(r_t^*)^2 x''(r_t^*)}{x'(r_t^*)^2} \frac{\partial r_t^*}{\partial f_d} = 0. \tag{B.60}$$

Then,

$$\begin{aligned} &\gamma^\delta \cdot \inf_r \left|\frac{2x(r)x'(r)^2 - x(r)^2 x''}{x'(r)^2}\right| \cdot \mathbb{E}_{\theta_0}\left(\left|\frac{\partial r_t^*(\theta_0, D_t^*)}{\partial f_d}\right| \mid \mathcal{F}_t\right) && \text{(B.61)} \\ &\le \mathbb{E}_{\theta_0}\left(f_{D_t^*} \cdot \left|\frac{2x(r_t^*)x'(r_t^*)^2 - x(r_t^*)^2 x''(r_t^*)}{x'(r_t^*)^2} \cdot \frac{\partial r_t^*(\theta_0, D_t^*)}{\partial f_d}\right| \mid \mathcal{F}_t\right) && \text{(B.62)} \\ &= \mathbb{E}_{\theta_0}\left(\left|\frac{\partial^2 V(D_t^*, T-t)}{\partial f_d \partial t}/m_0 + \frac{x(r_t^*(\theta_0, D_t^*))^2}{x'(r_t^*(\theta_0, D_t^*))}\right| \mid \mathcal{F}_t\right) && \text{(B.63)} \\ &= \mathbb{E}_{\theta}\left(\left|\frac{\partial^2 V(D_t^*, T-t)}{\partial f_d \partial t}/m_0 + \frac{\partial V(D_t^*, T-t)}{\partial t}/(f_{D_t^*} m_0)\right| \mid \mathcal{F}_t\right) && \text{(B.64)} \end{aligned}$$

where (B.62) follows from (B.59), (B.63) follows from (B.60), (B.64) is from (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 (B.61) is $2\sup_r |rx(r)| \le 2C_x$ from Assumption 1(v). Moreover, also from Assumption 1, we can show that $\inf_r \left|\frac{2x(r)x'(r)^2 - x(r)^2 x''(r)}{x'(r)^2}\right| \ge \frac{C_d \bar{x}^l}{M^2}$, then

$$\mathbb{E}_{\theta_0}\left(\left|\frac{\partial r_t^*(\theta_0, D_t^*)}{\partial f_d}\right| \mid \mathcal{F}_t\right) \le \frac{2C_x M^2}{\gamma^\delta C_d \bar{x}^l}. \tag{B.65}$$

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 (B.58) and (B.65), we have

$$\mathbb{E}_{\theta_0}\left[|r_t^*(\theta_0, D_t^*) - r_t^*(\theta_0, D_t^\pi)| \mid \mathcal{F}_t\right] \le \frac{2C_x M^2}{\gamma^\delta C_d \bar{x}^l} \cdot (p_0 + q_0) \cdot \mathbb{E}_{\theta_0}\left[\left|\frac{D_t^\pi}{m_0} - \frac{D_t^*}{m_0}\right| \mid \mathcal{F}_t\right]$$

Therefore, it follows that

$$\begin{aligned}
&\mathbb{E}\left[(r_t^*(\theta_0, D_t^*) - r_t^*(\theta_0, D_t^\pi))^2 \mid \mathcal{F}_t\right] = \mathcal{O}\left(\mathbb{E}_{\theta_0}\left[\left(\frac{D_t^*}{m_0} - \frac{D_t^\pi}{m_0}\right)^2 \mid \mathcal{F}_t\right]\right) \\
&\le \alpha\left(\frac{t}{e^t}\right)^2 \mathbb{E}\left[(r_t^*(\theta_0, D_t^*) - r_t^*(\theta_0, D_t^\pi))^2 \mid \mathcal{F}_t\right] + \mathcal{O}\left(\frac{1}{m_0}\right)
\end{aligned} \tag{B.66}$$

for some $\alpha > 0$ independent of m_0 and t, where the last relationship follows due to Lemma EC.2. This concludes our bound on (B).

The first term in right-hand side of (B.66) does not scale up with the problem scale m_0, so it is dominated by $\mathcal{O}(1/m_0)$. Thus, this proves the lemma. ∎

EC.2.13. Proof of Theorem 2

The Bayesian Cramer-Rao bound will be useful in our proof of Theorem 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 EC.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 of Theorem 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 regret. Hence, we assume only $\theta_0=q_0$ is unknown.

First, using the Bayesian Cramer-Rao inequality (Lemma EC.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 EC.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 $\tilde{\pi}\in\Pi$,*

$$\mathbb{E}_\theta\left[(r_t^{\tilde{\pi}}-r_t^*)^2\mid\mathcal{F}_t\right]\geq\alpha\left(\frac{1}{D_t^{\tilde{\pi}}}\right)\tag{B.67}$$

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

Proof of Claim EC.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 EC.4, the density function given $\mathcal{F}_t$ is

$$f(X\mid\theta)=\prod_{i=0}^{D_t}f_i(\theta),$$

where $f_i(\theta)$, $i=0,1,\dots,D_t$ are defined in (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 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 (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

$$\left(\mathbb{E}_\theta\left[\frac{\mathrm{d}}{\mathrm{d}\theta}g(\theta)\right]\right)^2 = \left(\mathbb{E}_\theta\left[\frac{\mathrm{d}}{\mathrm{d}\theta}r_t^*(\theta)\right]\right)^2 \geq \Omega(1). \tag{B.68}$$

Also, we have

$$\mathbb{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\mu(\theta)\right)^2\right] = \mathbb{E}_\theta\left[16\pi^2\left(\cos(\pi(\theta-3/4))\sin(\pi(\theta-3/4))\right)^2\right] \leq 16\pi^2. \tag{B.69}$$

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):

$$\begin{aligned}\mathbb{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln f(X\mid\theta)\right)^2\mid\mathcal{F}_t\right] &= -\mathbb{E}_\theta\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\ln f(X\mid\theta)\mid\mathcal{F}_t\right] = -\mathbb{E}_\theta\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\sum_{i=0}^{D_t}\ln f_i(\theta)\mid\mathcal{F}_t\right]\\ &= \sum_{i=0}^{D_t}-\mathbb{E}_\theta\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\ln f_i(\theta)\mid\mathcal{F}_t\right].\end{aligned}$$

Thus,

$$\mathbb{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\ln f(X\mid\theta)\right)^2\mid\mathcal{F}_t\right] = \sum_{i=0}^{D_t}-\mathbb{E}_\theta\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\ln f_i(\theta)\mid\mathcal{F}_t\right] \leq \sum_{i=0}^{D_t}\frac{1}{1\cdot {q_0}^2} = (D_t+1)/{q_0}^2. \tag{B.70}$$

Hence, taking (B.68),(B.69),(B.70) into Lemma EC.4, we have

$$\mathbb{E}_\theta\left[(r_t^{\tilde\pi}-r_t^*)^2\mid\mathcal{F}_t\right] = \mathbb{E}_\theta\left[(\hat g(X)-g(\theta))^2\mid\mathcal{F}_t\right] \geq \alpha_3\left(\frac{1}{D_t+1+16\pi^2}\right) = \alpha\left(\frac{1}{D_t+1}\right) \tag{B.71}$$

for some $\alpha > 0, \alpha_3 > 0$ independent of m_0.

end proof of Claim EC.3 ■

Next, we want to apply Claim EC.2 to prove the lower bound on regret. Thus, we need to check whether Assumption 2 and the condition of Proposition 3 hold. Notice that, $\left|\frac{\partial^2}{\partial r^2}[re^{-r}]\right| = (2-r)e^{-r}$, so Assumption 2 holds. Also, from Claim EC.3, we know $\mathbb{E}\left[(r_t^{\tilde\pi}-r_t^*)^2\mid\mathcal{F}_t\right] \geq \alpha\left(\frac{1}{D_t^{\tilde\pi}+1}\right)$. Since

$$\mathbb{E}\left(\frac{1}{D_t^{\tilde\pi}+1}\mid\mathcal{F}_t\right) = \frac{1}{\int_0^t\left(m_0-D_s^{\tilde\pi}\right)\left(p_0+q_0\frac{D_s^{\tilde\pi}}{m_0}\right)\mathrm{d}s+1} \geq \frac{1}{\bar x^u m_0(p_0+q_0)t} \geq \alpha' te^{-t}/m_0,$$

for some α' independent of t and m_0. This implies that the condition of Proposition 3 is satisfied. Hence, from Claim EC.2, we have

$$\begin{aligned}R^* - R(\tilde\pi) &= \Omega\left(\mathbb{E}\left[\int_0^T\xi(D_t^*)(r_t^{\tilde\pi}-r_t^*)^2\mathrm{d}t\right]\right)\\ &= \Omega\left(\mathbb{E}\left[\int_0^T\xi(D_t^{\tilde\pi})(r_t^{\tilde\pi}-r_t^*)^2\mathrm{d}t\right]\right) &\text{(B.72)}\\ &= \Omega\left(\mathbb{E}\left[\int_0^T\xi(D_t^{\tilde\pi})\frac{1}{D_t^{\tilde\pi}+1}\mathrm{d}t\right]\right) &\text{(B.73)}\end{aligned}$$

where (B.72) comes from the same analysis of (B.42) by replacing $\xi(D_t^*)$ by $\frac{\xi(D_t^*)}{\xi(D_t^{\tilde{\pi}})}\xi(D_t^{\tilde{\pi}})$, and (B.73) comes from Claim EC.3.

Then, we finally prove the lower bound on regret. Since $\mathbb{E}(D_t^{\tilde{\pi}}+1 \mid \mathcal{F}_t) \le \max_{s\le t}\xi(D_s^{\tilde{\pi}})\cdot 1\cdot(t+t_0)$ with $t_0=\Theta\left(m_0^{-1}\right)$ (we add t_0 here to avoid meaningless cases where $t=0$), we know

$$\text{(B.73)} \ge \alpha\left(\mathbb{E}\left[\int_0^T \frac{\xi(D_t^{\tilde{\pi}})}{\max_{s\le t}\xi(D_s^{\tilde{\pi}})}\frac{1}{t+t_0}\mathrm{d}t\right]\right) \ge \alpha\left(\mathbb{E}\left[\int_0^T \frac{\gamma^\delta}{(p_0+q_0)^2/(4q_0)}\frac{1}{t+t_0}\mathrm{d}t\right]\right) = \Omega(\ln m_0) \tag{B.74}$$

for some $\alpha>0$ independent of m_0, where the second inequality comes from (B.59).

This concludes our proof. ■

EC.2.14. Proof of Theorem 3

Proof. To simplify notation in this proof, we refer to $\hat{\theta}_0$ as θ instead. Let $D^\pi=(D_t^\pi, t\ge 0)$ 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 1) under parameter set θ and state (d,t).

Note that $R^*-R(\pi)$ is the regret of mis-specifying the demand parameter as θ, when the true parameter is θ_0. Since Assumption 2 holds, then according to Claim EC.2, we have

$$\begin{aligned}
R^*-R(\pi) &\ge \mathbb{E}_{\theta_0}\left[\int_0^T \xi(D_t^*)(r_t^\pi-r_t^*)^2\mathrm{d}t\right]\\
&\ge m_0\left[\min_{F\in[0,1]}(1-F)(p_0+q_0F)\right]\cdot\mathbb{E}_{\theta_0}\left[\int_0^T \mathbb{E}_{\theta_0}\left[(r_t^\pi-r_t^*)^2\mid\mathcal{F}_t\right]\mathrm{d}t\right]\\
&\ge m_0\left[\min_{F\in[0,1]}(1-F)(p_0+q_0F)\right]\cdot\int_0^T\Theta\left(\left(\frac{p-p_0}{p_0}\right)^2+\left(\frac{q-q_0}{q_0}\right)^2+\left(\frac{m-m_0}{m_0}\right)^2\right)\mathrm{d}t\\
&=\mathcal{E}^2 T\Omega(m_0).
\end{aligned}$$

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 4. The equality is due to $\left(\frac{p-p_0}{p_0}\right)^2+\left(\frac{q-q_0}{q_0}\right)^2+\left(\frac{m-m_0}{m_0}\right)^2=\mathcal{E}^2$. ■

EC.2.15. Proof of Theorem 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 $r^\pi=(r_t^\pi,t\ge 0)$ and $D^\pi=(D_t^\pi,t\ge 0)$, respectively. The price process under the oracle policy is $r^*=(r_t^*,t\ge 0)$.

We will use Proposition 3 and Lemma 4 to prove the theorem. Therefore, we need to check whether the conditions required in Proposition 3 and Lemma 4 are satisfied.

The condition required in Lemma 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 the policy applies MLE only when $\widehat{D}_t \geq 3$, 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 4.

Since the price for the first three customers is fixed, the regret in period $[0, t_3]$ is $\mathcal{O}(1)$. This is because the expected time of the third adoption is $\mathbb{E}(t_3) = \Theta(3/m_0)$. Following from Proposition 3, the regret for the first three adoptions is upper bounded by $\mathcal{O}\left(\frac{3}{m_0}\frac{3+1}{1/m_0}\right) = \mathcal{O}(1)$. Hence, to prove $R^* - R(\pi) = \mathcal{O}(\ln m_0)$, according to Proposition 3, it suffices to show that

$$\mathbb{E}\left[\int_{t_3}^{T} \frac{D_t^\pi + 1}{t + t_0}(r_t^\pi - r_t^*)^2 \mathrm{d}t\right] = \mathcal{O}(\ln m_0).$$

We know from Lemma 3 that, for any $t \in (t_3, T]$, the conditional expected estimation errors are

$$\mathbb{E}\left(\left(\frac{\hat{p}_t - p_0}{p_0}\right)^2 + \left(\frac{\hat{q}_t - q_0}{q_0}\right)^2 + \left(\frac{\hat{m}_t - m_0}{m_0}\right)^2 \mid \mathcal{F}_t\right) = \mathcal{O}\left(\frac{1}{D_t^\pi + 1}\right), \tag{B.75}$$

Also, we know $\mathcal{O}(1/(D_t^\pi + 1))$ dominates $\mathcal{O}(1/m_0)$ since $D_t^\pi \leq m_0$. Hence, by Lemma 4, it suffices to show that

$$\mathbb{E}\left[\int_{t_3}^{T} \frac{D_t^\pi + 1}{t + t_0}\left[\left(\frac{p_0 - \hat{p}_t}{p_0}\right)^2 + \left(\frac{q_0 - \hat{q}_t}{q_0}\right)^2 + \left(\frac{m_0 - \hat{m}_t}{m_0}\right)^2\right] \mathrm{d}t\right] = \mathcal{O}(\ln m_0).$$

By conditioning on $\mathcal{F}_t$,

$$\begin{aligned} &\mathbb{E}\left(\mathbb{E}\left[\int_{t_3}^{T} \frac{D_t^\pi + 1}{t + t_0}\left|\frac{\hat{p}_t - p_0}{p_0}\right|^2 \mathrm{d}t \mid \mathcal{F}_T\right]\right) + \mathbb{E}\left(\mathbb{E}\left[\int_{t_3}^{T} \frac{D_t^\pi + 1}{t + t_0}\left|\frac{\hat{q}_t - q_0}{q_0}\right|^2 \mathrm{d}t \mid \mathcal{F}_T\right]\right) \\ &+ \mathbb{E}\left(\mathbb{E}\left[\int_{t_3}^{T} \frac{D_t^\pi + 1}{t + t_0}\left|\frac{\hat{m}_t - m_0}{m_0}\right|^2 \mathrm{d}t \mid \mathcal{F}_T\right]\right) \leq \mathbb{E}\left(\int_{0}^{T} \frac{1}{t + t_0}\mathrm{d}t\right) = \mathcal{O}(\ln(m_0)), \end{aligned}$$

where the last inequality follows from (B.75). This proves the theorem. ■

EC.2.16. Proof of Theorem 5

Proof. For convenience, we will use π to denote the MBP-MLE-Limited policy $\pi^{\mathsf{M\text{-}Lim}}$. 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 $\hat{\theta}^\pi = (\hat{\theta}_t^\pi, t \geq t_{C_0})$ 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 oracle policy is $D^* = (D_t^*, t \geq 0)$.

Since the price for the first C_0 customers is fixed, and C_0 is independent of m_0, the regret in period $[0, t_{C_0}]$ is $\mathcal{O}(1)$. Hence, according to Proposition 3, it suffices to examine the bound for

$$\mathbb{E}\left[\mathbb{E}\left(\int_{t_{C_0}}^{T} \frac{D_t^\pi + 1}{t + t_0} (r_t^* - r_t^\pi)^2 \, \mathrm{d}t \mid \mathcal{F}_T\right)\right]. \tag{B.76}$$

With probability 1, we can decompose the pricing error at any $t \in (t_{C_0}, T]$ 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 4 (Section EC.2.12), specifically from (B.66), taking the expectation of (A) conditioning on $\mathcal{F}_t$ is bounded by

$$\frac{1}{m_0} = \mathcal{O}\left(\mathbb{E}\left[\left(\frac{p_0 - \hat{p}_t^\pi}{p_0}\right)^2 + \left(\frac{q_0 - \hat{q}_t^\pi}{q_0}\right)^2 + \left(\frac{m_0 - \hat{m}_t^\pi}{m_0}\right)^2 \mid \mathcal{F}_t\right]\right),$$

where the equality is from Lemma 3. According to Lemma 4, (C) is also bounded by $\mathcal{O}\left(\mathbb{E}\left[\left(\frac{p_0 - \hat{p}_t^\pi}{p_0}\right)^2 + \left(\frac{q_0 - \hat{q}_t^\pi}{q_0}\right)^2 + \left(\frac{m_0 - \hat{m}_t^\pi}{m_0}\right)^2 |\mathcal{F}_t\right]\right)$.

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 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 regret 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 regret due to (B) is zero since

$$\begin{aligned}
&\sum_{i=1}^{c_t} r^*_{t+\tau_{i-1}} \xi(D_t^\pi + i - 1) x(r^*_{t+\tau_{i-1}}) \Delta t_i - \sum_{i=1}^{c_t} r_t^\pi \xi(D_t^\pi + i - 1) x(r_t^\pi) \Delta t_i \\
&= \sum_{i=1}^{c_t} r_i \xi(D_t^\pi + i - 1) x(r_i) \Delta t_i - \left(\sum_{i=1}^{c_t} \xi(D_t^\pi + i - 1) \Delta t_i\right) \left(\frac{\sum_{j=1}^{c_t} r_j x(r_j) \xi(D_t^\pi + j - 1) \Delta t_j}{\sum_{j=1}^{c_t} \xi(D_t^\pi + j - 1) \Delta t_j}\right) = 0.
\end{aligned}$$

where the first equality is from (4.7). However, regret 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 (B.76), to prove the theorem, it suffices to bound

$$\mathbb{E}\left[\int_{t_{C_0}}^{T} \frac{D^{\pi}_t+1}{t+t_0}\left[(p_0-\hat{p}^{\pi}_t)^2+(q_0-\hat{q}^{\pi}_t)^2+\frac{1}{{m_0}^2}(m_0-\hat{m}^{\pi}_t)^2\right]\mathrm{d}t\big|\mathcal{F}_T\right]. \tag{B.77}$$

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 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 3, we know that on any demand sample path, (B.77) can be bounded above by

$$\begin{aligned}
&\int_{t_{C_{[0]}}}^{t_{C_{[0]}+1}} \frac{C_{[0]}+1}{t+t_0}\frac{1}{C_{[0]}+1}\mathrm{d}t+\int_{t_{C_{[0]}+1}}^{t_{C_{[0]}+2}} \frac{C_{[0]}+1+1}{t+t_0}\frac{1}{C_{[0]}+1}\mathrm{d}t+\cdots+\int_{t_{C_{[0]}+C_1-1}}^{t_{C_{[0]}+C_1}} \frac{C_{[0]}+C_1-1+1}{t+t_0}\frac{1}{C_{[0]}+1}\mathrm{d}t\\
&+\cdots\\
&+\int_{t_{C_{[K-1]}}}^{t_{C_{[K-1]}+1}} \frac{C_{[K-1]}+1}{t+t_0}\frac{1}{C_{[K-1]}+1}\mathrm{d}t+\cdots+\int_{t_{C_{[K-1]}+C_K-1}}^{t_{C_{[K-1]}+C_K}} \frac{C_{[K-1]}+C_K-1+1}{t+t_0}\frac{1}{C_{[K-1]}+1}\mathrm{d}t\\
&\le \int_{t_{C_{[0]}}}^{t_{C_{[1]}}} \frac{C_{[1]}}{t+t_0}\frac{1}{C_{[0]}+1}\mathrm{d}t+\int_{t_{C_{[1]}}}^{t_{C_{[2]}}} \frac{C_{[2]}}{t+t_0}\frac{1}{C_{[1]}+1}\mathrm{d}t+\ldots+\int_{t_{C_{[K-1]}}}^{t_{C_{[K]}}} \frac{C_{[K]}}{t+t_0}\frac{1}{C_{[K-1]}+1}\mathrm{d}t\\
&\le \left(1+\max_{i=1,2,\ldots,K}\frac{C_i}{C_{i-1}}\right)\int_0^T \frac{1}{t+t_0}\mathrm{d}t=\mathcal{O}\left(\left(1+\max_{i=1,2,\ldots,K}\frac{C_i}{C_{i-1}}\right)\ln(m_0T)\right),
\end{aligned}$$

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

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

EC.2.17. Proof of Proposition 4

Proof. We only need to analyze the concavity of the first term in (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

$$\mathbf{z}^\top\nabla^2_{\mu'}\phi(\mu')\mathbf{z} = -\sum_{i=0}^{\widehat{D}_t-1}\frac{(\mathbf{z}^\top y^{i,t_{i+1}})^2}{(\mu'^\top y^{i,t_{i+1}})^2}\le 0. \tag{B.78}$$

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.

■

EC.2.18. Proof of Lemma 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 5.4, note that the ML estimator $\hat\mu_t=(\hat\beta_t,\hat\gamma_t)$ is unique since $D_t\ge 3(n+1)$ and $\mathbf{Y}$ is full rank. Note from (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 $\bar{\bar{\delta}}_j, j=0,1,\cdots,n$ such that $\gamma_0-\bar{\bar{\delta}}_j\le\hat\gamma_j\le\gamma_0+\bar{\bar{\delta}}_j$.

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:

$$\hat\mu_t=\arg\max_{\substack{\mu:\beta_1\ge 0,\\ \beta_3\le 0}}\mathcal{L}_t(\widehat{\mathbf{U}}_t;\mu)=\mu_0+\arg\min_{\substack{u:u_{b1}\ge-\beta_{01},\\ u_{b3}\le-\beta_{03}}}-\sum_{i=0}^{D_t}\ln\frac{f_i(\mu_0+u)}{f_i(\mu_0)},$$

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, $\hat\mu_t=\mu_0+\hat u$.

To complete the proof of Lemma 3′, we will need to establish that $\mathbb{E}_{\mu_0}[|\hat\gamma_{tj}-\gamma_{0j}|^2\mid D_t=k]\le\frac{\alpha_{\gamma_j}}{k+1}$ for all $j=0,\dots,n$ and for some α_{γ_j} independent of m_0.

We examine the estimation error $|\hat\gamma_{tj}-\gamma_{0j}|$ for some $j=0,\dots,n$. 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 $\bar{\bar{\delta}}_j\ge\delta>0$. 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

$$-\sum_{i=0}^{D_t}\ln\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}\le 0\quad\text{ or }\quad -\sum_{i=0}^{D_t}\ln\frac{f_i(\mu_0-\delta e_j)}{f_i(\mu_0)}\le 0.$$

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 $i=0,\dots,D_t$.

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,

$$\begin{aligned}
&\mathbb{P}_{\mu_0}\{|\hat{\gamma}_{tj}-\gamma_{0j}|>\delta\} \\
&\le \mathbb{P}_{\mu_0}\left\{-\sum_{i=0}^{D_t}\ln\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}\le 0\right\}+\mathbb{P}_{\mu_0}\left\{-\sum_{i=0}^{D_t}\ln\frac{f_i(\mu_0-\delta e_j)}{f_i(\mu_0)}\le 0\right\} \\
&=\mathbb{P}_{\mu_0}\left\{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}\ge 1\right\}+\mathbb{P}_{\mu_0}\left\{\prod_{i=0}^{D_t}\frac{f_i(\mu_0-\delta e_j)}{f_i(\mu_0)}\ge 1\right\} \\
&\le \mathbb{P}_{\mu_0}\left\{\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\ge 1\right\}+\mathbb{P}_{\mu_0}\left\{\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0-\delta e_j)}{f_i(\mu_0)}}\ge 1\right\} \\
&\le \mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\right)+\mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0-\delta e_j)}{f_i(\mu_0)}}\right).
\end{aligned} \tag{B.79}$$

Hence, we need to bound the two terms in (B.79). 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 (B.79) can be written as

$$\mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\right)=\mathbb{E}_{\mu_0}\left(\cdots\mathbb{E}_{\mu_0}\left(\mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)\mid\mathcal{F}_{t_{D_t-2}}\right)\cdots\mid\mathcal{F}_0\right). \tag{B.80}$$

We will analyze this expression, starting from the innermost conditional expectation.

Note that

$$\begin{aligned}
&\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right)=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\mathbb{E}_{\mu_0}\left(\sqrt{\frac{f_{D_t}(\mu_0+\delta e_j)}{f_{D_t}(\mu_0)}}\mid\mathcal{F}_{t_{D_t-1}}\right) \\
&=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\left(\int_{t_{D_t-1}}^{\infty}\sqrt{\frac{f_{D_t}(\mu_0+\delta e_j)}{f_{D_t}(\mu_0)}}f_{D_t}(\mu_0)\mathrm{d}t_{D_t}\right) \\
&=\sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\left(\int_{t_{D_t-1}}^{\infty}\sqrt{f_{D_t}(\mu_0+\delta e_j)}\sqrt{f_{D_t}(\mu_0)}\mathrm{d}t_{D_t}\right).
\end{aligned} \tag{B.81}$$

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,

$$\int_{t_{D_{t-1}}}^{\infty}\sqrt{f_{D_t}(\mu_0+\delta e_j)f_{D_t}(\mu_0)}\mathrm{d}t_{D_t}=1-\frac{1}{2}\int_{t_{D_{t-1}}}^{\infty}\left(\sqrt{f_{D_t}(\mu_0+\delta e_j)}-\sqrt{f_{D_t}(\mu_0)}\right)^2 \mathrm{d}t_{D_t}. \tag{B.82}$$

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,

$$\frac{1}{2}\int_{t_{D_{t-1}}}^{\infty}\left(\sqrt{f_{D_t}(\mu_0+\delta e_j)}-\sqrt{f_{D_t}(\mu_0)}\right)^2 \geq \frac{1}{4\sqrt{R}}\mathbb{E}_{\mu_0}\left(\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\mu_0+\delta e_j)}\mid \mathcal{F}_{t_{D_{t-1}}}\right), \tag{B.83}$$

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 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

$$\begin{aligned}
&\mathbb{E}_{\mu_0}\left(\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\mu_0+\delta e_j)}\mid\mathcal{F}_{t_{D_{t-1}}}\right)=\mathbb{E}_{\mu_0}\left(\int_0^\delta\frac{\partial}{\partial z}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\mu_0+ze_j)}\mathrm{d}z\mid\mathcal{F}_{t_{D_{t-1}}}\right)\\
&=\mathbb{E}_{\mu_0}\left(\int_0^\delta\left(\frac{\partial}{\partial z}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\theta_0+ze_j)}-\frac{\partial}{\partial z}\ln\frac{f_{D_t}(\mu_0)}{f_{D_t}(\theta_0+ze_j)}\Big|_{z=0}\right)\mathrm{d}z\mid\mathcal{F}_{t_{D_{t-1}}}\right)\\
&=\mathbb{E}_{\mu_0}\left(\int_0^\delta\int_0^z\frac{\partial^2}{\partial z'^2}\ln\frac{f_{D_t}(\theta_0)}{f_{D_t}(\mu_0+z'e_j)}\mathrm{d}z'\mid\mathcal{F}_{t_{D_{t-1}}}\right)\geq\frac{1}{2C_I}\delta^2.
\end{aligned}$$

Then, (B.83) reduces to

$$\frac{1}{4\sqrt{R}C_I}\delta^2 \le \int_{t_{D_t-1}}^{\infty} \left(\sqrt{f_{D_t}(\mu_0+\delta e_j)} - \sqrt{f_{D_t}(\mu_0)}\right)^2 \mathrm{d}t_{D_t-1}.$$

Hence, from (B.82), we have

$$\begin{aligned}\int_{t_{D_t-1}}^{\infty} \sqrt{f_{D_t}(\mu_0+\delta e_j) f_{D_t}(\mu_0)}\mathrm{d}t_{D_t} &= 1 - \frac{1}{2}\int_{t_{D_t-1}}^{\infty} \left(\sqrt{f_{D_t}(\mu_0+\delta e_j)} - \sqrt{f_{D_t}(\mu_0)}\right)^2 \mathrm{d}t_{D_t} \\ &\le \exp\left(-\frac{1}{2}\int_{t_{D_t-1}}^{\infty} \left(\sqrt{f_{D_t}(\mu_0+\delta e_j)} - \sqrt{f_{D_t}(\mu_0)}\right)^2 \mathrm{d}t_{D_t}\right) \le \exp\left(-\frac{1}{8\sqrt{R}C_I}\delta^2\right),\end{aligned}$$

where the first inequality is because $e^{-x} \ge 1-x$ for all x.

Hence, from (B.81), we have

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}} \mid \mathcal{F}_{t_{D_t-1}}\right) \le \sqrt{\prod_{i=0}^{D_t-1}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\exp\left(-\frac{1}{8\sqrt{R}C_I}\delta^2\right). \tag{B.84}$$

This provides a bound for the innermost conditional expectation in (B.80). Observe that all of the terms in the right-hand side of (B.84) 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 (B.84) given $\mathcal{F}_{t_{D_t-2}}$, and use the same logic as our arguments above, we get

$$\mathbb{E}_{\mu_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}} \mid \mathcal{F}_{t_{D_t-2}}\right) \le \sqrt{\prod_{i=0}^{D_t-2}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\cdot\exp\left(-\frac{2}{8\sqrt{R}C_I}\delta^2\right)$$

We can proceed iteratively to evaluate (B.81) as we take conditional expectations given $\mathcal{F}_{t_{D_t}-3}$, $\mathcal{F}_{t_{D_t}-4}$, $\mathcal{F}_0$, resulting in

$$\mathbb{E}_{\theta_0}\left(\sqrt{\prod_{i=0}^{D_t}\frac{f_i(\mu_0+\delta e_j)}{f_i(\mu_0)}}\right) \le \mathbb{E}_{\mu_0}\left(\exp\left(-\frac{D_t+1}{8\sqrt{R}C_I}\delta^2\right)\right)$$

Using similar arguments, we can get the same bound for the second term in (B.79). Therefore, we have

$$\mathbb{P}_{\mu_0}\{|\hat{\gamma}_{tj}-\gamma_{0j}| > \delta\} \le 2\mathbb{E}_{\mu_0}\left(\exp\left(-\frac{D_t+1}{8\sqrt{R}C_I}\delta^2\right)\right).$$

Hence,

$$\begin{aligned}\mathbb{E}_{\mu_0}\left[(\hat{\gamma}_{tj}-\gamma_{0j})^2 \mid D_t = k\right] &= \int_0^{\infty} \mathbb{P}_{\mu_0}\left\{(\hat{\gamma}_{tj}-\gamma_{0j})^2 > \delta \mid D_t = k\right\}\mathrm{d}\delta \\ &= \int_0^{\infty} \mathbb{P}_{\mu_0}\left\{|\hat{\gamma}_{tj}-\gamma_{0j}|^2 > \sqrt{\delta} \mid D_t = k\right\}\mathrm{d}\delta \\ &\le \int_0^{\infty} \exp\left(-\frac{k+1}{8\sqrt{R}C_I}\delta\right)\mathrm{d}\delta = \frac{8\sqrt{R}C_I}{k+1}.\end{aligned}$$

■

EC.2.19. Proof of Lemma 4′

Proof. The proof of Lemma 4′ follows exactly the same steps as the proof of Lemma 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 (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,

$$\left|\frac{\partial r_t^*(\mu,d)}{\partial \gamma_i}\right| = \left|\frac{\frac{\partial}{\partial \gamma_i}[V(d,T-t)-V(d+1,T-t)] - \frac{b_{i,n}(r_t^*)x'(r_t^*) - x(r_t^*)\binom{n}{i}{r_t^*}^{i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{x'(r_t^*)^2}}{\frac{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}{x'(r_t^*)^2}}\right|. \tag{B.85}$$

Similar to (B.56), we have

$$\begin{aligned}\underbrace{\frac{\partial}{\partial \gamma_i}[V(d,T-t)-V(d+1,T-t)]}_{\text{(A1)}} &= \underbrace{\frac{\partial^2 V(d,T-t)}{\partial \gamma_i \partial t}\cdot\frac{1}{(m-d)\left(p+\frac{d}{m}q\right)x(r_t^*)}}_{\text{(A2)}} \\ &\quad - \frac{2x(r_t^*)x'(r_t^*)b_{i,n}(r_t^*) - x(r_t^*)^2\binom{n}{i}{r_t^*}^{i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{x'(r_t^*)^2}.\end{aligned} \tag{B.86}$$

We now examine the absolute value of (A2) on the right-hand side of (B.86). 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:

$$|\text{(A1)}| \le \frac{r_t^* b_{i,n}(r_t^*)}{x(r_t^*)} + \left|\frac{2x(r_t^*)x'(r_t^*)b_{i,n}(r_t^*) - x(r_t^*)^2\binom{n}{i}{r_t^*}^{i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{x'(r_t^*)^2}\right|. \tag{B.87}$$

Therefore, we substitute (B.87) into (B.85) to get

$$\begin{aligned}\left|\frac{\partial r_t^*(\mu,d)}{\partial \gamma_i}\right| &\le \left|\frac{r_t^* b_{i,n}(r_t^*)x'(r_t^*)^2}{x(r_t^*)\left(2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)\right)}\right| + \left|\frac{2x(r_t^*)x'(r_t^*)b_{i,n}(r_t^*) - x(r_t^*)^2\binom{n}{i}{r_t^*}^{i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right| \\ &\quad + \left|\frac{b_{i,n}(r_t^*)x'(r_t^*) - x(r_t^*)\binom{n}{i}{r_t^*}^{i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{2x'(r_t^*)^2 - x(r_t^*)x''(r_t^*)}\right| \\ &\le \frac{M^2}{\bar{x}^l C_d}b_{i,n}(r_t^*) + \frac{2M\bar{x}^u}{C_d}b_{i,n}(r_t^*) + \left|\frac{\bar{x}^{u2}\binom{n}{i}{r_t^*}^{i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{C_d}\right| + \frac{M}{C_d}b_{i,n}(r_t^*) \\ &\quad + \left|\frac{\bar{x}^u\binom{n}{i}{r_t^*}^{i-1}(1-r_t^*)^{n-i-1}(i-nr_t^*)}{C_d}\right|\end{aligned}$$

where the inequality follows from Assumption 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.

$\blacksquare$