INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

WILEY

INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

Intl. Trans. in Op. Res. 21 (2014) 737–760
DOI: 10.1111/itor.12064

# A simulation-based algorithm for optimal pricing policy under demand uncertainty

Saswata Chakravarty[a], Sindhu Padakandla[b] and Shalabh Bhatnagar[b]

[a] *Department of Electrical Engineering, Indian Institute of Science, Bangalore 560012, India*
[b] *Department of Computer Science and Automation, Indian Institute of Science, Bangalore 560012, India*
*E-mail: saswata.123@gmail.com [Chakravarty]; sindhupr@csa.iisc.ernet.in [Padakandla]; shalabh@csa.iisc.ernet.in [Bhatnagar]*

Received 27 May 2013; accepted 30 October 2013

**Abstract**

We propose a simulation-based algorithm for computing the optimal pricing policy for a product under uncertain demand dynamics. We consider a parameterized stochastic differential equation (SDE) model for the uncertain demand dynamics of the product over the planning horizon. In particular, we consider a dynamic model that is an extension of the Bass model. The performance of our algorithm is compared to that of a myopic pricing policy and is shown to give better results. Two significant advantages with our algorithm are as follows: (a) it does not require information on the system model parameters if the SDE system state is known via either a simulation device or real data, and (b) as it works efficiently even for high-dimensional parameters, it uses the efficient smoothed functional gradient estimator.

*Keywords:* optimal pricing policy; Bass model; parameterized stochastic differential equation (SDE); stochastic approximation algorithm; smoothed functional gradient estimates

## 1. Introduction

A new product introduced in the market undergoes dynamic changes in demand during its life cycle in the market. Diffusion and market saturation effects are two notable causes for such changes. The diffusion effect leads to an increase in the rate of sales with increase in market penetration due to word-of-mouth communication among buyers and other factors such as advertising. The saturation effect, on the other hand, is the decrease in the rate of sales with increase in market penetration when the market potential is finite. Dynamics may also be present in the supply side. For instance, it could result in a decrease in the unit cost of production with increase in cumulative production.

Pricing of products in the presence of such demand and supply dynamics is a difficult but important decision problem for the producer. While keeping a high price may increase profits for the short term, in the long run, the market penetration may be affected and the net profit over


Published by John Wiley & Sons Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main St, Malden, MA02148, USA.

the life cycle of the product may decrease. Thus, there is a need to formulate a mathematical optimization problem to compute the optimal pricing policy or rule in order to maximize the net profit over the planning horizon. Optimal pricing in models incorporating the effect of market dynamics has been studied in the literature (Araman and Caldentey, 2009; Chinthalapati et al., 2006; Farias and Van Roy, 2010; Lin, 2006). These models assume that the market demand is known with certainty over the planning horizon and follows a deterministic path (which may be controlled by the price). However, such an assumption is quite unrealistic. A stochastic extension of the optimal pricing model, which explicitly considers the uncertainty in demand, was proposed in Raman and Chatterjee (1995), where the cumulative sales of the product is modeled as a stochastic process governed by a stochastic differential equation (SDE) parametrized by the price policy.

In this paper, we propose a simulation-based method to compute the optimal price policy of the stochastic optimization model discussed in Raman and Chatterjee (1995). We consider a parameterized SDE whose drift term is obtained from the generalized Bass model. We perform a time discretization on this SDE by using an Euler–Milstein discretization scheme (Glasserman, 2003). The problem then becomes one of finding an optimal parameter trajectory over a finite time horizon, for which we adapt an algorithm from Bhatnagar et al. (2009) (see also Chapter 14 of Bhatnagar et al., 2013). While the algorithm of Bhatnagar et al. (2009) applies for the case where the drift integrates over time (in the associated integral equation), our adaptation of the same works for the case where the drift integrates over the sample path of the state trajectory. Upon convergence, our algorithm gives an optimal price trajectory over the entire time horizon. The algorithm incorporates the efficient smoothed functional gradient estimates of Bhatnagar (2007) and requires only two system simulations at each iteration while updating the entire parameter trajectory. We prove the convergence of our algorithm. We show the results of several experiments with different cost and system parameters. We also derive a myopic policy using the model presented in Raman and Chatterjee (1995) and observe that our algorithm consistently shows better results as compared to the myopic policy. A significant advantage with our algorithm is that, unlike model of Raman and Chatterjee (1995), it does not require a priori knowledge of the system parameters, as long as information on the SDE state is made available either through a simulation device or real data. Thus, one may use any model, not necessarily the generalized Bass model, when using our algorithm. Also, our algorithm efficiently updates a 400-epoch parameter trajectory (that can alternatively be viewed in the setting of simulation optimization as a 400-dimensional parameter) at each instant because of the use of the smoothed functional gradient estimates. To the best of our knowledge, in the setting of simulation optimization, experiments on such high-dimensional settings involving expected finite or infinite-horizon cost structures have not been previously reported because most simulation optimization schemes do not work on such (high-dimensional) settings.

The rest of the paper is organized as follows. In Section 2, we briefly review the existing literature on the Bass model, which is a popular model of demand dynamics, as well as some of its extensions to optimal pricing models. In Section 3, we describe the problem formulation. In Section 4, we present our proposed simulation-based algorithm whose detailed convergence proof is given in Section 5. In Section 6, we show the experimental results using our algorithm on a few different problem settings and compare these with the myopic price policy. In Section 7, we present our concluding remarks and discuss future work. The derivation of the myopic price policy using the model in Raman and Chatterjee (1995) is given in the Appendix.

## 2. The Bass model and some extensions

Product diffusion models are used to model the level of spread of a new product (also commonly referred as innovation in the marketing literature) among a given set of prospective buyers or adopters over time. It is used to quantify how the number of adopters grows with time.

One of the most popular diffusion models in the marketing literature is the Bass model, proposed by F.M. Bass (Bass, 1969). It is a first purchase diffusion model that assumes that in the time horizon under consideration, there are no repeat buyers and purchase volume per buyer is one unit. Other popular models are those proposed by Fourt and Woodlock (1960) and Mansfield (1961).

### 2.1. The Bass model

The Bass model assumes that there are two driving forces that are primarily responsible for the growth of adoption of a new product, namely, mass-media communication and word-of-mouth communication (Mahajan et al., 1990). It also distinguishes between the adopters. In particular, those adopters who are influenced by mass-media communication are termed as "innovators," while those influenced by word-of-mouth communication are termed as "imitators." The Bass model describes the evolution of the cumulative number of adopters by time $t$ via an ODE. This is obtained by considering the following functional relationship between $f(t)$, the density function of the time to adoption, and $F(t) = \int_0^t f(t)dt$, the cumulative fraction of adopters by time $t$:

$$f(t)/(1 - F(t)) = p + qF(t). \tag{1}$$

The parameter $p$, called the "coefficient of innovation," reflects the effect of mass-media communication and the parameter $q$, called the "coefficient of imitation," reflects the effect of word-of-mouth communication.

Let $M$ denote the market potential, that is, the number of ultimate adopters. Then we can write $MF(t) = X(t)$, where $X(t)$ is the cumulative number of adopters by time $t$. Now, we can rewrite (1) as

$$f(t) = (1 - F(t))(p + qF(t)).$$

$$\Rightarrow \frac{dF(t)}{dt} = \left(1 - \frac{X(t)}{M}\right)\left(p + \frac{q}{M}X(t)\right).$$

$$\Rightarrow \frac{dX(t)}{dt} = (M - X(t))\left(p + \frac{q}{M}X(t)\right).$$

Thus, we get the "Bass model" as

$$\frac{dX(t)}{dt} = p(M - X(t)) + \frac{q}{M}X(t)(M - X(t)). \tag{2}$$

The first term in the differential Equation (2) represents adoptions due to buyers who are not influenced in the timing of their adoption by the number of people who have already bought the

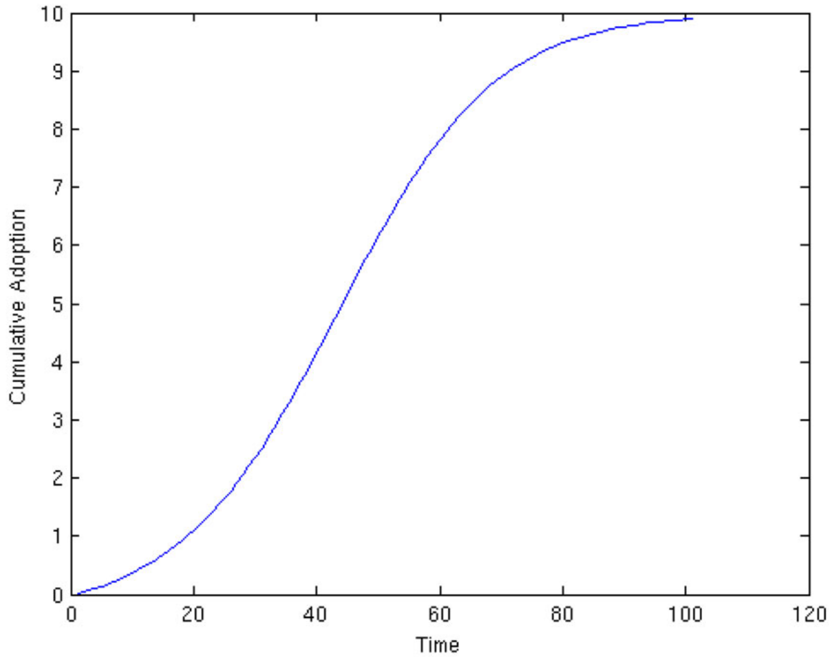



Fig. 1. Bass curve.

product (Mahajan et al., 1990), whereas the second term represents the adoptions that are due to buyers who are influenced by the number of previous buyers.

Figure 1 shows the plot of a typical solution to the Bass equation. It is an S-shaped curve, which well characterizes the diffusion and saturation effects.

### 2.2. *Extending the Bass model*

The basic Bass model given by Equation (2) has been extended in various directions. Equation (2) assumes that the market potential is constant over time. There have been models suggested by Kalish (1985), Mahajan and Peterson (1978), Jain and Rao (1990), and many others, where the market potential is treated as a function of control variables such as price or noncontrol variables such as growth in population and number of households.

The basic Bass model depicts the growth of a single product that is assumed independent of the growth of other products, but there are situations where this is not true. For example, consider accessories of a mobile phone (primary product) such as a Bluetooth headset. The growth of such accessories and their market potential depends upon the diffusion of the primary product. In other words, the growth dynamics of these products is coupled. This sort of coupling also occurs while modeling the growth of different generations of the same product.

The Bass model is essentially a demand model, as it characterizes the manner in which the demand of a product evolves without any consideration to supply restrictions. When there is insufficient supply to satisfy the demand, a queue of buyers would result and the distribution of the time to adopt would be similar to the supply distribution, as people can only adopt when supply becomes available. Such a model that incorporates supply restrictions has been suggested by Jain et al. (1991).

The Bass model in (2) does not consider the influence of marketing strategies. There have been several extensions to the Bass model proposed in the literature, which aim at incorporating the effect of marketing strategies such as price, advertising, and promotion. In some papers (e.g., Kalish, 1985) the market potential parameter of the Bass model has been expressed as a function of price. Other papers, such as Jain and Rao (1990), report that the price influences the coefficients of innovation and imitation $p$ and $q$, respectively. In Bass et al. (1994) and Robinson and Lakhani (1975), the effect of price on the adoption rate is assumed to be multiplicatively separable.

### 2.3. Optimal pricing

The extension of the Bass model incorporating price provides a setting for studying optimal strategies for product pricing. The first such work in the literature is by Robinson and Lakhani (1975), where the effect of price on demand is assumed to be multiplicatively separable. The adoption model considered in their article is as follows:

$$\frac{dX(t)}{dt} = (M - X(t))\left(p + \frac{q}{M}X(t)\right)e^{-\epsilon p_r(t)}, \tag{3}$$

where $p_r(t)$ is the price parameter at time $t$ and $\epsilon > 0$ is a constant.

In Krishnan et al. (1999), the adoption model uses a generalized Bass model of the following form:

$$\frac{dX(t)}{dt} = (M - X(t))\left(p + \frac{q}{M}X(t)\right)\left(1 + \beta\frac{\dot{p}_r(t)}{p_r(t)}\right). \tag{4}$$

Given the model of how the cumulative number of adopters evolve, controlled by the price parameter $p_r(t)$, it is desirable to obtain the optimal price path that maximizes the discounted profits over a given length of horizon $T$:

$$J = \int_0^T (p_r(t) - c)e^{-rt}dX(t), \tag{5}$$

where $c$ is the cost of production that may be constant or a function of $X$ and $r$ is the discount rate.

In Krishnan et al. (1999), the authors have tried to characterize the nature of the optimal price path. Although no closed-form solution for the optimal price path was developed in the general case, they have identified sets of conditions that influence the optimal policy. Their main result is that the optimal policy either monotonically decreases with time or shows an increase–decrease behavior. It depends on the price sensitivity parameter $\beta$ and the discount rate $r$, and is independent of the Bass model parameters (coefficients of innovation ($p$) and imitation ($q$) as well as the market potential ($M$)).

### 2.4. Optimal pricing under uncertainty

The Bass model or its extensions assume that the demand or the adoption of a product follows a deterministic path (which may be controlled by the price). However, such an assumption is quite unrealistic. There are many uncertainties involved in this process. Uncertainties in demand may be present due to uncertainties in the acceptance of products, as well as in response to advertisements and in gauging the market potential. So the deterministic Bass model parameters are insufficient to capture the demand dynamics of the market.

In Raman and Chatterjee (1995), the authors have studied the effect of demand uncertainty on the optimal price path. They consider a controlled SDE

$$dX(t) = b(X(t), p_r(t))dt + \sigma(X(t))dW(t). \tag{6}$$

Here the drift term, $b(X, p_r)$, is the demand model that characterizes the adoption rate and the diffusion term, $\sigma(X(t))$, characterizes the uncertainty in demand. Also, $W(t)$, $t \geq 0$ is the standard Brownian motion. The objective is to find the price path $p_r(t)$, $t \geq 0$ that maximizes the expected discounted cumulative profits over an infinite time horizon:

$$J = \max_{p_r(t)} E\left[\int_0^\infty e^{-rt}(p_r(t) - c(X(t)))dX(t)\right]. \tag{7}$$

The cost of production is $c(X(t))$. Three kinds of demand models are considered in the paper—a static model ($b(X, p_r) = p - qp_r$), Bass model with multiplicatively separable effect of price ($b(X, p_r) = (M - X)(p + \frac{q}{M}X)(1 - \gamma p_r)$), and a simple price-timing model ($b(X, p_r) = \alpha(M - \gamma p_r - X)$). For the first and third scenarios, a closed-form analytical expression was derived for the optimum policy. For the case of the Bass model with constant demand uncertainty and cost, only a polynomial approximation was developed, which is accurate only at low values of $X(t)$. The main result for the Bass model is that with greater demand uncertainty, the initial optimal price $p_r(0)$ increases and the magnitude of the initial slope of the price path $dp_r/dX$ decreases.

In the next two sections, we adapt a simulation-based multitimescale stochastic optimization algorithm based on the smoothed functional technique from Bhatnagar et al. (2009) to the problem of estimating the optimal price parameters. We consider the Bass model with multiplicatively separable price for the drift term in (6), namely the second demand model considered in Raman and Chatterjee (1995). We let the planning horizon be finite and do not consider discounting even though the same can be easily included as well.

## 3. Problem formulation

We are interested in finding the optimal price path for a new product over a finite planning horizon. The cumulative sales or adoption among the population is governed by the following SDE:

$$dX(t) = (M - X(t))\left(p + \frac{q}{M}X(t)\right)(1 - \gamma p_r(t))dt + \sigma(X(t))dW(t), \tag{8}$$

where $X(t)$ is the cumulative number of adopters or cumulative sales by time $t$, $M$ is the market potential, $p$ is the coefficient of innovation/external influence, $q$ is the coefficient of imitation/word-of-mouth influence, $\gamma$ is the parameter that determines the influence of price $p_r(t)$ on sales and $W(t)$, $t \geq 0$ is a Brownian motion process.

We are interested in finding the function $p_r^*(t), 0 \leq t \leq 1$ that minimizes over all $p_r(t)$ the expected difference between the expected long-term cost and profit over a horizon $T$:

$$J = \min_{p_r(t), 0 \leq t \leq T} E\left[\int_0^T (c(X(t)) - p_r(t))dX(t)\right]. \tag{9}$$

Note that the above will be minimized by minimizing the expected long-term cost and maximizing at the same time the expected long-term profit. For computational simplicity and to apply a simulation-based optimization algorithm in this setting, we first consider a discrete version of this problem by using time discretization. The time horizon $T$ is discretized into $N$ stages, each of duration $h$, that is, $T = Nh$. We denote $X_j \equiv X(jh)$ and $p_{r(j)} \equiv p_r(jh)$. Equation (8) is discretized according to the Euler–Milstein scheme (Glasserman, 2003) as follows:

$$X_{j+1} = X_j + b(X_j, p_{r(j)})h + \sigma(X_j)\sqrt{h}Z_{j+1} + \frac{1}{2}\sigma'(X_j)\sigma(X_j)h\big(Z_{j+1}^2 - 1\big). \tag{10}$$

Here $\sigma'(\cdot)$ is the derivative of $\sigma(\cdot)$, $b(X_j, p_{r(j)}) = (M - X_j)(p + \frac{q}{M}X_j)(1 - \gamma p_{r(j)})$ and $Z_j$, $1 \leq j \leq N$ are independent normally distributed samples with zero mean and unit standard deviation. The objective in the discrete domain now is to find the (discrete) parameter trajectory, that is, the set of parameters $\mathbf{p_r} = \{p_{r(0)}, p_{r(1)}, \ldots, p_{r(N-1)}\}$ that minimize

$$J_{X_0}(p_{r(0)}, p_{r(1)}, \ldots, p_{r(N-1)}) = \mathbb{E}\left[\sum_{j=0}^{N-1} g_j(X_j, p_{r(j)})(X_j - X_{j-1}) \,\middle|\, X_0\right] h,$$

where $g_j(X_j, p_{r(j)}) = (c(X_j) - p_{r(j)})$, and $X_{-1} = 0$. The initial state $X_0$ is assumed to be known. Since $h$ is a small time element that does not play a role in the minimization, we drop it and the objective in this case simply becomes

$$J_{X_0}(p_{r(0)}, p_{r(1)}, \ldots, p_{r(N-1)}) = \mathbb{E}\left[\sum_{j=0}^{N-1} g_j(X_j, p_{r(j)})(X_j - X_{j-1}) \,\middle|\, X_0\right]. \tag{11}$$

We apply an algorithm as described in Bhatnagar et al. (2009) for the solution. However, it must be noted that there is a difference in the objective function in Bhatnagar et al. (2009) and the one above. In Bhatnagar et al. (2009), the cost is integrated over time, whereas in (9), the cost is integrated over the path of the trajectory. Thus, the discretized objective is the weighted sum of the single stage costs.

The cost (11) can be decomposed into the costs per stage as

$$\begin{aligned} J_{X_0}(p_{r(0)}, p_{r(1)}, \ldots, p_{r(N-1)}) &= J_{X_0}^0(p_{r(0)}) + J_{X_0}^1(p_{r(0)}, p_{r(1)}) \\ &\quad + \cdots + J_{X_0}^{N-1}(p_{r(0)}, \ldots, p_{r(N-1)}), \end{aligned} \tag{12}$$

where $J^j_{X_0}(p_{r(0)}, \ldots, p_{r(j-1)})$ is the cost of the $j$th stage. Thus the cost-to-go at stage $j$ depends on the parameters $p_{r(0)}, \ldots, p_{r(j-1)}$, or alternatively, the parameter $p_{r(i)}$ affects the cost-to-go from each of the subsequent stages $i+1, \ldots, N$.

## 4. Two-timescale stochastic optimization algorithm with smoothed functional gradient estimates

We now describe the two-timescale stochastic optimization algorithm. We require two timescales in the update rule because our objective function is an average sum over single stage costs. Along the faster timescale, the gradient of the objective function is estimated, and along the slower timescale the price parameters are updated using a gradient estimate based on the smoothed functional technique.

### *4.1. Gradient estimation*

Simulation optimization is a technique for finding optimal parameters when the objective function involves an expectation and is not analytically tractable. In such a scenario, gradient search methods are useful. We incorporate the following two-simulation smoothed functional gradient estimate given in Bhatnagar (2007):

$$D_\beta J_{X_0}(\mathbf{p_r}) = \frac{1}{\beta} E\left[\frac{\eta}{2}(J_{X_0}(\mathbf{p_r} + \beta\eta) - J_{X_0}(\mathbf{p_r} - \beta\eta)) \mid \mathbf{p_r}\right], \tag{13}$$

where $D_\beta J_{X_0}(\mathbf{p_r})$ represents the gradient of the objective convolved with an $N$-dimensional multivariate Gaussian density, $\eta$ is a vector of independent $N(0, 1)$ distributed random variables, and $\beta > 0$ is a small constant. It can be shown (see Section 5) that $\| D_\beta J_{X_0}(\mathbf{p_r}) - \nabla J_{X_0}(\mathbf{p_r}) \| \to 0$ as $\beta \to 0$.

From the decomposition of the cost in (12), it is seen that $p_{r0}$ influences the cost for all stages, $p_{r1}$ influences the cost from the second stage onwards, and so on. Hence the partial derivative of the cost with respect to each of the parameters can be written as

$$\frac{\partial J_{X_0}(\mathbf{p_r})}{\partial p_{r(j)}} = \sum_{i=j}^{N-1} \frac{\partial J^i_{X_0}(p_{r(0)}, \ldots, p_{r(i-1)})}{\partial p_{r(j)}}. \tag{14}$$

### *4.2. The algorithm procedure*

Let $\{a(n)\}$ and $\{c(n)\}$ be two step-size sequences that satisfy the standard requirements of stochastic approximation, see Assumption 3. Let $\mathbf{p_r(n)} \triangleq (p_{r(j)}(n), j = 0, 1, \ldots, N-1)^\top$ denote the $n$th update of the price vector $\mathbf{p_r}$. Let $\eta(\mathbf{n}) \triangleq (\eta_j(n), j = 0, 1, \ldots, N-1)^\top$ be a vector of independent $N(0, 1)$ random variables. We let the price parameters $p_{r(j)}(n), \forall j = 0, 1, \ldots, N-1, n \geq 0$ take values in a compact interval $C = [A, B]$ for $0 < A < B$). In the algorithm below, this is ensured through a projection operator $\Gamma : \mathcal{R} \to C$ defined by $\Gamma(x) = \min(B, \max(A, x))$. In general, one may also consider different price intervals $C_j = [A_j, B_j]$ for different time instants $j = 0, 1, \ldots, N-1$.

First, we perturb the parameter vectors as suggested by (13), namely, $\mathbf{p_r^+(n)} \equiv \mathbf{p_r(n)} + \beta\eta(\mathbf{n})$ and $\mathbf{p_r^-(n)} \equiv \mathbf{p_r(n)} - \beta\eta(\mathbf{n})$, respectively, and obtain two sets of parallel simulations of the SDE that are governed by these two parameter vectors. Then, on the faster timescale we estimate the gradient using (13), and on the slower timescale the parameters are updated in the negative gradient direction. Better convergence behavior is observed if the gradient estimation is done over an additional $R$ loops ($R > 1$), where the perturbed parameters are kept fixed and the trajectory is generated $R$ times. In principle $R$ can be 1; however, it is generally observed that $R > 1$ with an arbitrarily chosen value of $R$ gives better performance.

The detailed algorithm is as follows.

**Algorithm 1.** Two-timescale smoothed functional algorithm

---

**Input:** $N$, the number of stages; $h$, the discretization step
**Input:** $R$: arbitrarily chosen integer, $L$: a large integer
**Input:** Model parameters: $M$, the market potential; $p$, the coefficient of innovation; $q$, the coefficient of imitation; $\gamma$, the price sensitivity factor; $X_0$, the initial number of adopters. Define $b(X_j, p_{r(j)}) = (M - X_j)(p + \frac{q}{M}X_j)(1 - \gamma p_{r(j)})$
**Input:** Noise function ($\sigma(X)$) parameter: $\sigma_0$.
**Input:** Algorithm parameters: $\mathbf{p_r}(0)$, initial price trajectory, $\beta$,$R$.

---

$n \leftarrow 0$
**loop**
Generate $\eta_j(n) \sim N(0, 1)$ *independently* $\quad 0 \le j \le N-1$
Obtain perturbed parameters:
$\mathbf{p_r^+} : p^+_{r(j)}(n) = p_{r(j)}(n) + \beta\eta_j(n), \quad 0 \le j \le N-1$
$\mathbf{p_r^-} : p^-_{r(j)}(n) = p_{r(j)}(n) - \beta\eta_j(n), \quad 0 \le j \le N-1$
for $m = 0$ to $R - 1$
**loop**
Obtain perturbed state trajectories $\{X_j^+(nR+m)\}$ and $\{X_j^-(nR+m)\}$:
for $j = 0$ to $N - 1$
**loop**
$Z^+_{j+1}(nR+m) \sim N(0,1), \quad Z^-_{j+1}(nR+m) \sim N(0,1)$

$$X^+_{j+1}(nR+m) = X_j^+(nR+m) + b(X_j^+(nR+m), p^+_{r(j)}(n))h$$
$$+ \sigma(X_j^+(nR+m))\sqrt{h}Z^+_{j+1}(nR+m) + \tfrac{1}{2}\sigma'(X_j^+(nR+m))$$
$$\times\sigma(X_j^+(nR+m))h((Z^+_{j+1}(nR+m) - 1))^2$$

$$X^-_{j+1}(nR+m) = X_j^-(nR+m) + b(X_j^-(nR+m), p^-_{r(j)}(n))h$$
$$+ \sigma(X_j^-(nR+m))\sqrt{h}Z^-_{j+1}(nR+m) + \tfrac{1}{2}\sigma'(X_j^-(nR+m))$$
$$\sigma(X_j^-(nR+m))h((Z^-_{j+1}(nR+m) - 1))^2$$

**end loop**
Estimate gradient: $\forall j = 0, 1, \ldots, N-1$,
$Y_j(nR+m+1) = (1 - c(n))Y_j(nR+m)$

$$+ c(n)\left[\frac{\eta_j(n)}{2\beta}\left(\sum_{i=j}^{N}(g_i(X_i^+(nR+m), p_{r(i)}^+(n))(X_i^+(nR+m) - X_{i-1}^+(nR+m))\right.\right.$$

$$\left.\left. - g_i(X_i^-(nR+m), p_{r(i)}^-(n))(X_i^-(nR+m) - X_{i-1}^-(nR+m)))\right) - Y_j(nR+m)\right].$$

**end loop**

Update price parameters:

$p_{r(j)}(n+1) = \Gamma_j\left(p_{r(j)}(n) - a(n)Y_j((n+1)R)\right), \quad \forall j = 0, \ldots, N-1.$

$n \leftarrow n+1$

**if** $n = L$ **then**

Terminate with $\mathbf{p_r}(L)$.

**end if**

**end loop**

---

**Remark 1.** We describe below an online implementation of the above algorithm that will work with real data. Let $X_j(n)$, $j = 0, 1, \ldots, N-1$, $n \geq 0$ denote a single state trajectory obtained from the real data observations. We update $p_{r(j)}(n)$, $\forall j$, once every $2R$ instants of time in place of $R$, with $p_{r(j)}^+(n)$ and $p_{r(j)}^-(n)$, held fixed for $2R$ instants. The trajectory $X_j(n)$ is now governed by $p_{r(j)}^+(n)$ in odd cycles $mR \leq n < (m+1)R$, $m = 0, 2, 4$, etc., and by $p_{r(j)}^-(n)$ in even cycles $mR \leq n < (m+1)R$, $m = 1, 3, 5$, etc. The gradient estimate is similar to the one in the algorithm except that the most recent values of $X_j(n)$ when governed by $p_{r(j)}^+(n)$ (resp. $p_{r(j)}^-(n)$) are used in place of $X_j^+(n)$ (resp. $X_j^-(n)$). Finally, the price parameters are updated once every $2R$ instants but using the same update procedure as in the algorithm.

## 5. Convergence analysis

The convergence analysis proceeds in the usual manner by showing that the noise terms asymptotically vanish and that the iterates asymptotically track the stable fixed points of an associated ordinary differential equation. In the following, we let $\|\cdot\|$ denote the Euclidean norm. We make the following assumptions:

**Assumption 1.** *The cost function $c(\cdot)$ is Lipschitz continuous and the diffusion term $\sigma(\cdot)$ is continuous.*

**Assumption 2.** *The functions $J_{X_0}(p_{r(0)}, p_{r(1)}, \ldots, p_{r(N-1)})$ are continuously differentiable in $p_{r(0)}, p_{r(1)}, \ldots, p_{r(N-1)}$ and have bounded second derivatives.*

**Assumption 3.** *The step sizes $a(n), c(n), n \geq 0$ satisfy the requirements: $a(n), c(n) > 0$, $\forall n$. Further,*

$$\sum_{n=0}^{\infty} a(n) = \sum_{n=0}^{\infty} c(n) = \infty, \quad \sum_{n=0}^{\infty}(a(n)^2 + c(n)^2) < \infty,$$

$$a(n) = o(c(n)).$$

Assumptions 1–3 are standard assumptions that are routinely used in the analysis of stochastic optimization algorithms. The Lipschitz continuity on the cost $c(\cdot)$ in Assumption 1 ensures an at most linear growth in each stage. This is because

$$|c(X)| - |c(0)| \leq |c(X) - c(0)| \leq \kappa |X|,$$

where $\kappa > 0$ is the Lipschitz constant. Thus, $|c(X)| \leq \hat{\kappa}(1 + |X|)$, where $\hat{\kappa} = \max(|c(0)|, \kappa)$. Recall that the drift term has the form $b(X, p_r) = (M - X)(p + \frac{q}{M}X)(1 - \gamma p_r)$ (derived from the Bass model) and is continuous as well. While we do not prescribe a form for the diffusion term, we consider various instances of $\sigma(\cdot)$ in our experiments.

Assumption 2 is used to push through a Taylor series argument to show the unbiasedness of the gradient estimator in the limit as $\beta \to 0$. Assumption 3 is also a standard requirement on the step-size sequences. The first requirement (summing to infinity of the step sizes) ensures that the algorithm's trajectory when mapped on the timescale of the ODE does not exhibit premature convergence. The noise and gradient estimation error terms are both seen to vanish as a consequence of the second requirement in Assumption 3. Also, the last condition implies that $a(n)$ goes to zero at a rate faster than $c(n)$. Thus, increments involving $a(n)$ are uniformly smaller than those involving $c(n)$ from some $N_0$ onwards (i.e., $\forall n \geq N_0$). As a result, recursions governed by $c(n)$ converge faster than those governed by $a(n)$ even though the former recursions exhibit higher variance in their iterates unlike those governed by $c(n)$. The latter iterates exhibit slow but more graceful convergence, that is, with less oscillations. The timescale corresponding to $a(n)$ is therefore also referred as the slower timescale, while the one corresponding to $c(n)$ is the faster of the two timescales.

We have the following as an immediate consequence of Assumption 1.

**Lemma 1.** *The functions $g_j(\cdot, \cdot)$ are Lipschitz continuous.*

*Proof.* From Assumption 1, there exists a constant $\alpha > 0$ such that

$$|c(X_j) - c(Y_j)| \leq \alpha |X_j - Y_j|.$$

Now,

$$\begin{aligned} |g_j(X_j, p_{r(j)}) - g_j(Y_j, p_{w(j)})| &= |c(X_j) - p_{r(j)} - c(Y_j) + p_{w(j)}| \\ &\leq |c(X_j) - c(Y_j)| + |p_{r(j)} - p_{w(j)}| \\ &\leq \alpha |X_j - Y_j| + |p_{r(j)} - p_{w(j)}| \\ &\leq (\alpha + 1)(|X_j - Y_j| + |p_{r(j)} - p_{w(j)}|). \end{aligned} \tag{15}$$

Now note that

$$(|X_j - Y_j| + |p_{r(j)} - p_{w(j)}|)^2 \leq 2(|X_j - Y_j|^2 + |p_{r(j)} - p_{w(j)}|^2).$$

Thus,

$$|X_j - Y_j| + |p_{r(j)} - p_{w(j)}| \leq \sqrt{2} \parallel (X_j, p_{r(j)}) - (Y_j, p_{w(j)}) \parallel .$$

From (15), it now follows that

$$|g_j(X_j, p_{r(j)}) - g_j(Y_j, p_{w(j)})| \leq K_1 \parallel (X_j, p_{r(j)}) - (Y_j, p_{w(j)}) \parallel,$$

where $K_1 = \sqrt{2}(\alpha + 1)$. The claim follows. □

We first analyze below the faster recursions, that is, those governed by the step sizes $c(n), n \geq 0$. For simplicity, we show the analysis here for the case of $R = 1$. The general case follows with a few minor modifications. Let $\mathcal{F}(l) = \sigma(p_{r(j-1)}(n), \eta_{j-1}(n), X_j^+(n), X_j^-(n), n \leq l, j = 1, \ldots, N), l \geq 1$, denote a sequence of associated $\sigma$ fields.

Consider the sequence $\{M_j(p)\}$ (for given $j \in \{1, \ldots, N\}$) defined as follows:

$$
\begin{aligned}
M_j(p) = \sum_{n=1}^{p} c(n)\Bigg(\frac{\eta_j(n)}{2\beta}\Bigg(\sum_{i=j}^{N}(g_i(X_i^+(n), p_{r(i)}^+(n))(X_i^+(n) - X_{i-1}^+(n))\\
- g_i(X_i^-(n), p_{r(i)}^-(n))(X_i^-(n) - X_{i-1}^-(n)))\Bigg)\Bigg)\\
- E\Bigg[\frac{\eta_j(n)}{2\beta}\Bigg(\sum_{i=j}^{N}(g_i(X_i^+(n), p_{r(i)}^+(n))\big(X_i^+(n) - X_{i-1}^+(n)\big)\\
- g_i(X_i^-(n), p_{r(i)}^-(n))(X_i^-(n) - X_{i-1}^-(n)))\Bigg)\Bigg|\mathcal{F}(n-1)\Bigg].
\end{aligned}
$$

**Lemma 2.** *The sequences $\{M_j(p), \mathcal{F}(p)\}$, $j = 1, \ldots, N$ are almost surely convergent martingale sequences.*

*Proof.* For simplicity of notation in the proof, we suppress the index $n$ in the quantities $X_l^+(n), X_l^-(n)$, $p_{r(l)}^+(n)$, and $p_{r(l)}^-(n)$, and simply denote these as $X_l^+, X_l^-, p_{r(l)}^+$, and $p_{r(l)}^-$, respectively.

It is easy to see that $\{M_j(p), \mathcal{F}(p)\}$, $j = 1, \ldots, N$ are martingale sequences. We now verify that each of the $M_j(p), \forall j = 1, \ldots, N, p \geq 0$ is a square integrable random variable. Note that

$$
\begin{aligned}
E\big[M_j^2(p)\big] \leq \frac{R_{p,j}}{4\beta^2}\sum_{n=1}^{p} c(n)^2 E\Bigg[\eta_j^2(n)\sum_{l=j}^{N}(g_l^2(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+) + g_l^2(X_l^-, p_{r(l)}^-)(X_i^- - X_{i-1}^-))\\
+ E^2\Bigg[\eta_j(n)\sum_{l=j}^{N}(g_l(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+) - g_l(X_l^-, p_{r(l)}^-)(X_l^- - X_{l-1}^-)) \mid \mathcal{F}(n-1)\Bigg]\Bigg], \quad (16)
\end{aligned}
$$

for some constant $R_{p,j} > 0$ (that depends on $p$, $j$, and $N$). By the conditional Jensen's inequality, we have

$$
E^2\left[\eta_j(n)\sum_{l=j}^{N}(g_l(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+) - g_l(X_l^-, p_{r(l)}^-)(X_l^- - X_{l-1}^-))|\mathcal{F}(n-1)\right]
$$

$$\leq E\left[\eta_j^2(n)\left(\sum_{l=j}^{N}(g_l(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+) - g_l(X_l^-, p_{r(l)}^-)(X_l^- - X_{l-1}^-))\right)^2 \middle| \mathcal{F}(n-1)\right]$$

$$\leq K_j E\left[\eta_j^2(n)\sum_{l=j}^{N}(g_l^2(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+)^2 + g_l^2(X_l^-, p_{r(l)}^-)(X_l^- - X_{l-1}^-)^2)\middle|\mathcal{F}(n-1)\right]$$

for some constant $K_j > 0$ (that depends on $j$ and $N$). Hence from (16), for some $R_{p,j} > 0$,

$$E\big[M_j^2(p)\big] \leq \frac{R_{p,j}}{4\beta^2}\sum_{n=1}^{p} c(n)^2 E[\eta_j^2(n)\sum_{l=j}^{N}(g_l^2(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+)^2$$

$$+ g_l^2(X_l^-, p_{r(l)}^-)(X_l^- - X_{l-1}^-)^2)$$

$$+ E\left[\eta_j^2(n)\sum_{l=j}^{N}(g_l^2(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+)^2 + g_l^2(X_l^-, p_{r(l)}^-)(X_l^- - X_{l-1}^-)^2) \mid \mathcal{F}(n-1)]\right].$$

Thus,

$$E\big[M_j^2(p)\big] \leq \frac{R_{p,j}}{2\beta^2}\sum_{n=1}^{p} c(n)^2 E\Bigg[\eta_j^2(n)\sum_{l=j}^{N}(g_l^2(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+)^2$$

$$+ g_l^2(X_l^-, p_{r(l)}^-)(X_l^- - X_{l-1}^-)^2)\Bigg]$$

$$\leq \frac{R_{p,j}}{2\beta^2}\sum_{n=1}^{p} c(n)^2 \sum_{l=j}^{N} E\big[\eta_j^4(n)\big]^{1/2}\big(E\big[g_l^4(X_l^+, p_{r(l)}^+)(X_l^+ - X_{l-1}^+)^4)\big]^{1/2}$$

$$+ E\big[g_l^4(X_l^-, p_{r(l)}^-)(X_l^- - X_{l-1}^-)^4)\big]^{1/2}\big),$$

by the Cauchy–Schwartz inequality. Since, $g_l(\cdot,\cdot)$ are Lipschitz continuous functions, we have

$$|g_l(X_m, p_{r(m)})| - |g_l(0,0)| \leq |g_l(X_m, p_{r(m)}) - g_l(0,0)| \leq K_l \parallel (X_m, p_{r(m)}) \parallel$$

where $K_l > 0$ is the Lipschitz constant associated with $g_l(\cdot,\cdot)$. Thus,

$$|g_l(X_m, p_{r(m)})| \leq C_1(1+ \parallel (X_m, p_{r(m)} \parallel)$$

$$\leq C_1(1 + |X_m| + |p_{r(m)}|),$$

for $C_1 = \max(K_l, |g_l(0,0)|) < \infty$. Thus,

$$E\left[g_l^4(X_m, p_{r(m)})(X_m - X_{m-1})^4)\right] \leq C_2 E\left[\left(1 + X_m^4\right)\left(X_m^4 + X_{m-1}^4\right)\right],$$

for some constant $C_2 > 0$. It is easy to see that $\sup_n E[(X_l^+)^4]$, $\sup_n E[(X_l^+)^8] < \infty$ since $l \leq N < \infty$ for given $N$ and the only sources of randomness are the $N$ random variables $Z_l^+$, $l = 1, \ldots, N$ in any sample path, each of which is $N(0,1)$-distributed. Moreover, as stated before, the parameter $p_{r(l)}$, $l = 0, 1, \ldots, N-1$ takes values in the compact set $C$, and both functions $b(\cdot,\cdot)$ and $\sigma(\cdot)$ are continuous. Similarly, $\sup_n E[(X_l^-)^4]$, $\sup_n E[(X_l^-)^8] < \infty$ as well. Thus, $E[M_j^2(p)] < \infty$, for all $p \geq 1$. It is now easy to see from Assumption 3 (square summability of the step-size sequence $c(n)$) and the fact that $\sup_n E[(X_l^+)^4 + (X_l^+)^8 + (X_l^-)4 + (X_l^-)^8] < \infty$ that

$$\sum_p E[(M_j(p+1) - M_j(p))^2 \mid \mathcal{F}(p)] < \infty \text{ a.s.}$$

Thus, by the martingale convergence theorem (cf. Theorem 3.3.4, pp. 53–54 of Borkar (1995)), that $\{M_j(p)\}$ are almost surely convergent martingale sequences. □

Before proceeding further, we recall the following important result from Hirsch (1989) (cf. Theorem 1, pp. 339). Consider the following ODE in $\mathcal{R}^N$:

$$\dot{z}(t) = f(z), \tag{17}$$

for a Lipschitz continuous $f : \mathcal{R}^N \to \mathcal{R}^N$ such that (17) has a globally asymptotically stable attractor $\zeta$. Given $\epsilon > 0$, let $\zeta^\epsilon$ denote the $\epsilon$-neighborhood of $\zeta$. Given $T, \Delta > 0$, we call a bounded, measurable $x(\cdot) : \mathcal{R}^+ \cup \{0\} \to \mathcal{R}^N$, a $(T, \Delta)$-perturbation of (17) if there exist $0 = T_0 < T_1 < T_2 < \cdots < T_r \uparrow \infty$ with $T_{r+1} - T_r \geq T$ $\forall r$ and solutions $z^r(y)$, $y \in [T_r, T_{r+1}]$ of (17) for $r \geq 0$, such that $\sup_{y \in [T_r, T_{r+1}]} \| z^r(y) - x(y) \| < \Delta$. The following is the result from Hirsch (1989).

**Lemma 3.** *Given $\epsilon$, $T > 0$, $\exists \bar{\Delta} > 0$ such that for all $\Delta \in (0, \bar{\Delta})$, every $(T, \Delta)$-perturbation of (17) converges to $\zeta^\epsilon$.*

Let $Y(n) \stackrel{\triangle}{=} (Y_1(n), \ldots, Y_N(n))^T$. Consider now the following system of ODEs:

$$\dot{\mathbf{p}}_{\mathbf{r}}(t) = 0, \tag{18}$$

$$\dot{Y}(t) = D_{\beta,2} J_{X_0}(\mathbf{p}_{\mathbf{r}}(t)) - Y(t). \tag{19}$$

As a consequence of (18), the ODE (19) can be rewritten according to

$$\dot{Y}(t) = D_{\beta,2} J_{X_0}(\mathbf{p}_{\mathbf{r}}) - Y(t), \tag{20}$$

that is, $\mathbf{p}_{\mathbf{r}}(t) \equiv \mathbf{p}_{\mathbf{r}}$ $\forall t$. Define two sequences $\{s(n)\}$ and $\{t(n)\}$ as follows: $s(0) = t(0) = 0$, $s(n) = \sum_{i=0}^{n-1} a(i)$, $t(n) = \sum_{i=0}^{n-1} c(i)$, $n \geq 1$, respectively. Define $\bar{Y}(\cdot), \bar{\mathbf{p}}_{\mathbf{r}}(\cdot)$ as follows: $\bar{Y}(t(n)) = Y(n)$, $\bar{\mathbf{p}}_{\mathbf{r}}(t(n)) = \mathbf{p}_{\mathbf{r}}(n)$ $\forall n$ with a linear interpolation on the intervals $[t(n), t(n+1)]$.

**Lemma 4.** *We have* $\| Y(n) - D_{\beta,2}J_{X_0}(\mathbf{p_r}(n)) \| \to 0$ *with probability 1, as* $n \to \infty$.

*Proof.* Note that the slower ($\mathbf{p_r}$) recursion can be rewritten as

$$p_{r(j)}(n+1) = \Gamma_j\left(p_{r(j)}(n) - c(n)\epsilon(n)\right),$$

where $\epsilon(n) = \frac{a(n)}{c(n)} Y_j(n)$. Since $a(n) = o(c(n))$ (cf. Assumption 3), it follows that $\epsilon(n) = o(1)$. Consider now the ODE (20). Since $\sup_n E[(X_l^+)^4 + (X_l^+)^8 + (X_l^-)4 + (X_l^-)^8] < \infty$ (see proof of Lemma 2) and the fact that $\mathbf{p_r} \in C^N$ (a compact set), it is easy to see that $\sup_{X_0,\mathbf{p_r}} D_{\beta,2}J_{X_0}(\mathbf{p_r}) < \infty$. As a consequence of Lemma 2 and the fact that $c(n) \to 0$ as $n \to \infty$, it follows that beyond some $N_1$ (i.e., for all $n \geq N_1$), $Y_j(n+1)$ is a convex combination of $Y_j(n)$ and a uniformly bounded quantity. Hence, the $Y_j(n)$ iterates remain uniformly bounded almost surely. Thus, for any $T, \delta > 0$, $(\bar{Y}(t(n)+\cdot), \mathbf{p_r}(t(n)+\cdot))$ is a bounded $(T,\delta)$-perturbation of (18 and 19) for $n$ sufficiently large. The claim now follows by applying Lemma 3 for every $\epsilon > 0$. □

**Lemma 5.** $\| Y(n) - \nabla J_{X_0}(\mathbf{p_r}(n) \| \to 0$ *with probability 1, as* $n \to \infty$.

*Proof.* It can be shown as in Proposition A.3 of Bhatnagar et al. (2009) that

$$\| D_{\beta,2}J_{X_0}(\mathbf{p_r}(n)) - \nabla J_{X_0}(\mathbf{p_r}(n)) \| \to 0,$$

as $\beta \to 0$. The rest now follows from Lemma 4 by using the triangle inequality. □

We now analyze the slower recursion. We will use a key result from Kushner and Clark (cf. Theorem 5.3.1 of Kushner and Clark, 1978), also stated as Theorem E.1 of Bhatnagar et al. (2013). We first describe the result below. Consider the following $N$-dimensional stochastic recursion:

$$X_{n+1} = \Gamma(X_n + \alpha(n)(h(X_n) + \beta_n + \xi_n)), \tag{21}$$

where $\Gamma : \mathcal{R} \to C^N$ as before, under the conditions (B1)–(B5) listed below. Also, consider the following ODE associated with (21):

$$\dot{X} = \tilde{\Gamma}(h(X(t))), \tag{22}$$

where for any $y \in \mathcal{R}^N$ and a bounded, continuous function $v(\cdot) : \mathcal{R}^N \to \mathcal{R}^N$,

$$\tilde{\Gamma}(v(y)) = \lim_{0<\eta\to 0}\left(\frac{\Gamma(y+\eta v(y)) - \Gamma(y)}{\eta}\right).$$

The operator $\tilde{\Gamma}$ ensures that the ODE (23) evolves within the constraint set $C^N$. Let $\bar{K}$ denote the set of asymptotically stable fixed points of the ODE (22).

Let $\tau(n), n \geq 0$ be a sequence of positive real numbers defined according to $\tau(0) = 0$ and for $n \geq 1$, $\tau(n) = \sum_{j=0}^{n-1}\alpha(j)$. Let $m(\tau) = \max\{n \mid \tau(n) \leq \tau\}$. Thus, $m(\tau) \to \infty$ as $\tau \to \infty$.

Now consider the following set of conditions for (21).

(B1) The function $h : \mathcal{R}^N \to \mathcal{R}^N$ is continuous.

(B2) The step sizes $\alpha(n), n \geq 0$ satisfy

$$\alpha(n) > 0 \forall n, \ \sum_n \alpha(n) = \infty, \ \alpha(n) \to 0 \text{ as } n \to \infty.$$

(B3) The sequence $\beta_n, n \geq 0$ is a bounded random sequence with $\beta_n \to 0$ almost surely as $n \to \infty$.
(B4) There exists $T > 0$ such that $\forall \epsilon > 0$,

$$\lim_{n\to\infty} P\left(\sup_{j\geq n}\max_{t\leq T}\left|\sum_{i=m(jT)}^{m(jT+t)-1} a(i)\xi_i\right| \geq \epsilon\right) = 0.$$

(B5) The set $C^N$ in which $X_n$ take values is a compact subset of $\mathcal{R}^N$.

Theorem 5.3.1 of Kushner and Clark (1978) says the following:

**Theorem 1.** *Under Conditions (B1)–(B5), $X_n \to \bar{K}$ as $n \to \infty$ almost surely.*

We will show the convergence of the slower recursion by verifying (B1)–(B5) in our case. The ODE associated with the $Y(n)$ iterates is the following:

$$\dot{\mathbf{p}}_{\mathbf{r}} = \tilde{\Gamma}(-\nabla J_{X_0}(\mathbf{p}_{\mathbf{r}}(t))). \tag{23}$$

The set of stable fixed points of (23) lies within the set $K \stackrel{\triangle}{=} \{\mathbf{p}_{\mathbf{r}} \in C^N \mid \tilde{\Gamma}(-\nabla J_{X_0}(\mathbf{p}_{\mathbf{r}})) = 0\}$. Note that $J_{X_0}(\cdot)$ itself serves as the associated Liapunov function for (23). Given $\epsilon > 0$, let $K^\epsilon = \{\mathbf{p}_{\mathbf{r}} \in C^N \mid \| \mathbf{p}_{\mathbf{r}} - \mathbf{p}_{\mathbf{r}0} \| < \epsilon$, for some $\mathbf{p}_{\mathbf{r}0} \in K\}$. The following is our main convergence result.

**Theorem 2.** *Given $\epsilon > 0$, there exists a $\hat{\beta} > 0$, such that for all $\beta \in (0, \hat{\beta}]$, the sequence $\{\mathbf{p}_{\mathbf{r}}(n)\}$ obtained using our algorithm converges to a point in $(C^N)^\epsilon$ with probability 1 as $n \to \infty$.*

*Proof.* The proof will follow from Theorem 1 once we verify the conditions (B1)–(B5) for our setting. The correspondence of the slower recursion with (21) is the following: $\mathbf{p}_{\mathbf{r}}(n) \equiv X_n$, $a(n) \equiv \alpha(n)$, $\nabla J_{X_0}(\mathbf{p}_{\mathbf{r}}(n)) \equiv h(X_n)$, $\beta_n = Y(n) - \nabla J_{X_0}(\mathbf{p}_{\mathbf{r}}(n))$, and $\xi_n \equiv 0$. Now (B1) follows as a consequence of Assumption 2. Also, (B2) follows from Assumption 3. In particular, the step sizes $a(n) \to 0$ because $\sum_n a(n)^2 < \infty$. It has been shown in Lemma 4 that the iterates $Y(n)$ are uniformly bounded. Further, because of Assumption 2 and the fact that $\mathbf{p}_{\mathbf{r}}(n)$ take values in the compact set $C^N$, $\nabla J_{X_0}(\mathbf{p}_{\mathbf{r}}(n))$ is uniformly bounded on the set $C^N$ as well. Thus, $\beta_n, n \geq 0$ are uniformly bounded and tend to 0 from Lemma 5. Thus, (B3) follows. Condition (B4) is trivially satisfied since $\xi_n \equiv 0$ in our case. Finally, condition (B5) is satisfied from the definition of $C^N$ as a compact set. The claim now follows by applying Theorem 1. □

## 6. Simulation results

The optimal pricing algorithm for the product growth model (8) discussed in the previous section is now applied to different problem settings. For all cases, the algorithm parameters $\beta$ and $R$ are set to 1 and 25, respectively. The step-size sequences are chosen to be: $a(n) = \frac{1}{\lceil \frac{n}{1000} \rceil^{\frac{3}{4}}}$ and $c(n) = \frac{1}{\lceil \frac{n}{1000} \rceil^{\frac{2}{3}}}$, respectively. Here $\lceil \frac{n}{1000} \rceil$ denotes the (integer-valued) ceiling of the division of $n$ by 1000. For instance, $\lceil \frac{999}{1000} \rceil = 1$, while $\lceil \frac{1001}{1000} \rceil = 2$. Thus, the step sizes are fixed (constant) for 1000 runs of the algorithm after which their values are changed. This is seen to help in better convergence behavior. The algorithm is run in each case for a total of $L = 50{,}000$ iterations.

The values of the model parameters are chosen from the scenario presented in section 4.2 of Raman and Chatterjee (1995). In particular, the various parameter values selected are $p = 0.02$, $q = 0.5$, and $\gamma = 0.005$. The values of the coefficient of innovation ($p$) and imitation ($q$) are based on typical estimates of the Bass model in earlier studies. The initial state $X_0$ was set to 0. The time horizon $T$ is chosen to be 100 units. The discretization step $h = 0.25$ and thus the number of stages $N = 400$. The time horizon was chosen so as to allow the saturation effect in the Bass model to set in.

We now present the experimental results for three different scenarios—constant uncertainty and constant cost, cost learning with constant uncertainty, and decreasing uncertainty with cost learning.

### 6.1. Constant uncertainty and constant cost

In this setting, we assume that the diffusion term is constant, that is, $\sigma(X_j) = \sigma_0$. The cost of production $c(X_j)$ is set to 80. The market potential $M$ was set to 10. The value of $\sigma_0$ was varied between 0.1 and 0.8, and the optimal price trajectory was obtained in each case. Then the system was simulated with the optimal trajectory with 100 different initial seeds and the mean and standard error values of the objective function were computed.

The convergence plots for some of the parameters in the case of $\sigma_0 = 0.5$ are shown in Fig. 2. The results for the various values of $\sigma_0$ are shown in Figs. 3–6. Figure 3(a) shows how the cumulative number of customers evolve for two different price paths—the optimal and myopic prices. For this setting, the myopic pricing policy (see Appendix for a derivation) is calculated using Equation (34) in Raman and Chatterjee (1995) and is found to give a constant value of 140. Figure 3(b) shows the optimal price trajectory.

Figures 3–6 show similar plots for different values of $\sigma_0$. It is observed that the initial price in the optimal price path increases with an increase in $\sigma_0$, which substantiates the claim in Raman and Chatterjee (1995).

Table 1 shows the values of the objective function obtained for the optimal price trajectory and myopic price. The values shown are the 95% confidence interval of the mean values over 100 independent simulation runs. It is seen that the optimal price path obtained outperforms the myopic price setting.

### 6.2. Cost learning with constant uncertainty

In this setting, we again consider the uncertainty to be constant, however with a varying cost. In many situations, with increase in sales level, and therefore an increase in production, the cost per product comes down. In this setting, we assume the cost to decrease linearly with the sales level. The market potential $M$ was set here to 10. The function $c(X_j)$ is set to $100 - 0.2X_j$. As in the previous setting, the value of $\sigma_0$ was varied from 0.1 to 0.8 and the optimal price trajectory was obtained in each case. Then the system was simulated using the optimal trajectory with 100 different initial seeds and the mean and standard error values of the objective function were computed.

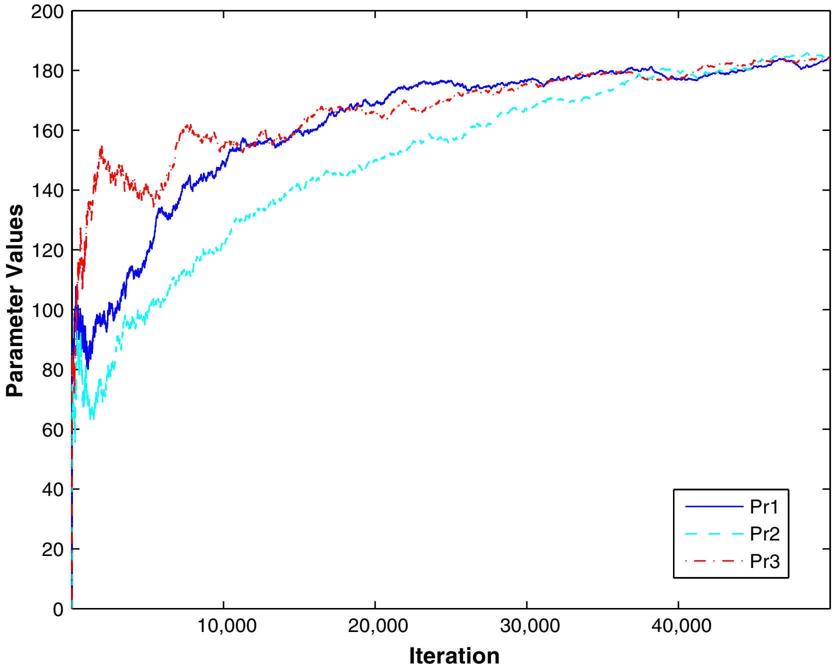
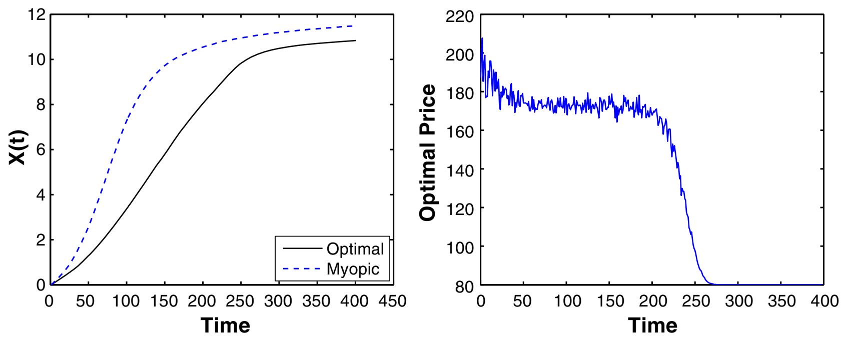


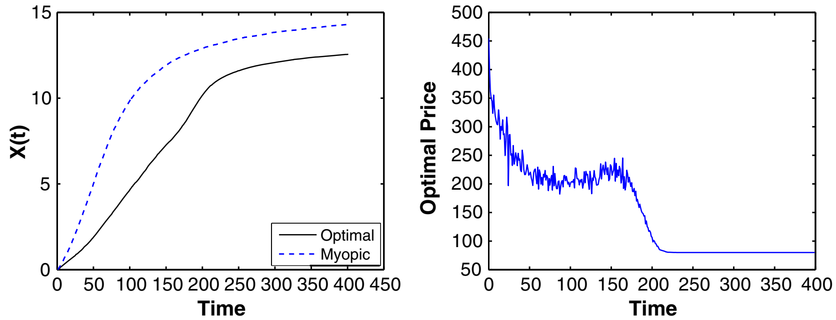

Fig. 2. Convergence plots for three parameter components.




**(a)** State trajectory **(b)** Optimal price path

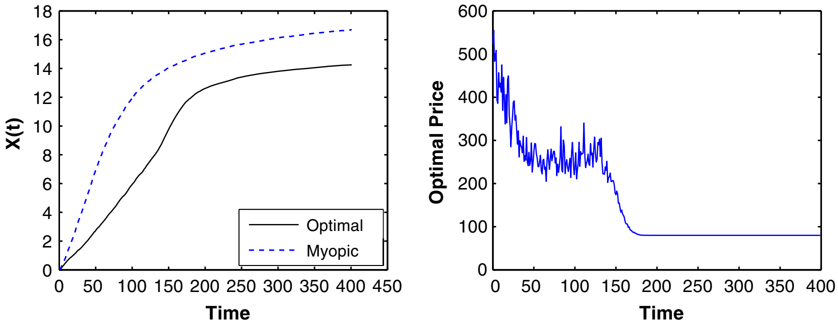
Fig. 3. $\sigma_0 = 0.1$.

Figures 7–10 show the results for different values of $\sigma_0$. The observations are similar to those of the first setting. From Table 2, it is seen that our algorithm outperforms the myopic policy. For higher values of $\sigma_0$, the difference between the objective function values obtained using the algorithm and the myopic policy is large.



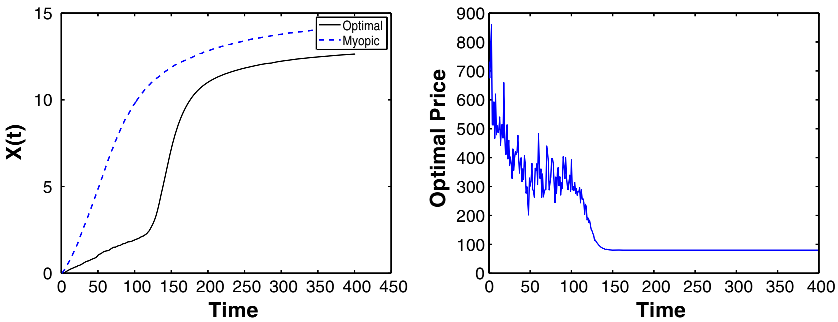

**(a)** State trajectory **(b)** Optimal price path

Fig. 4. $\sigma_0 = 0.3$.




**(a)** State trajectory **(b)** Optimal price path

Fig. 5. $\sigma_0 = 0.5$.

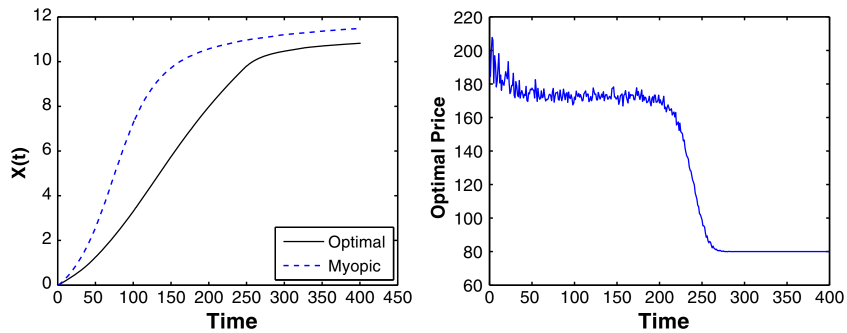



**(a)** State trajectory **(b)** Optimal price path

Fig. 6. $\sigma_0 = 0.8$.

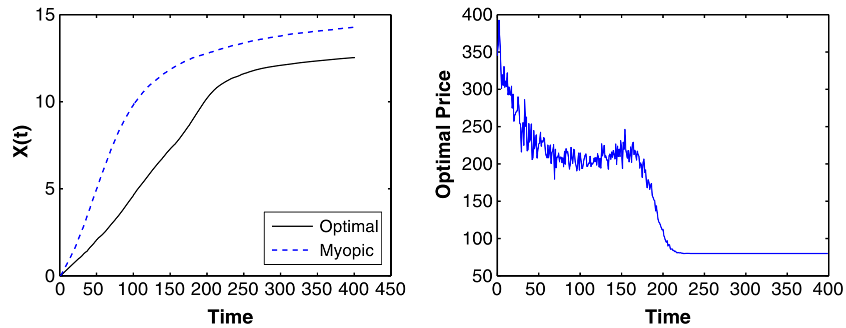
Table 1
Objective function values for the optimal and myopic prices with 95% confidence interval of mean values for the setting of constant uncertainty and constant cost

| $\sigma_0$ | Optimal price path | Myopic price |
|---|---|---|
| 0.1 | – 1028.159 ± 8.1438 | – 590.3236 ± 1.97764 |
| 0.3 | – 1106.389 ± 29.1452 | – 593.22 ± 2.10504 |
| 0.5 | – 780.84 ± 21.3836 | – 590.828 ± 2.26 |
| 0.8 | – 590.828 ± 2.26 | – 522.529 ± 16.58944 |




**(a)** State trajectory **(b)** Optimal price path

Fig. 7. $\sigma_0 = 0.1$.



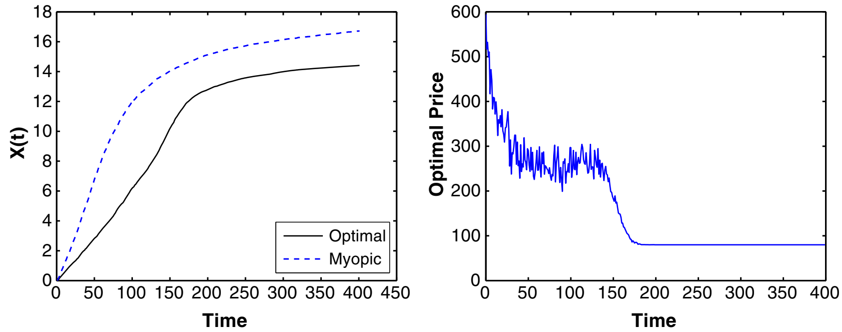

**(a)** State trajectory **(b)** Optimal price path

Fig. 8. $\sigma_0 = 0.3$.

### 6.3. Decreasing uncertainty with cost learning

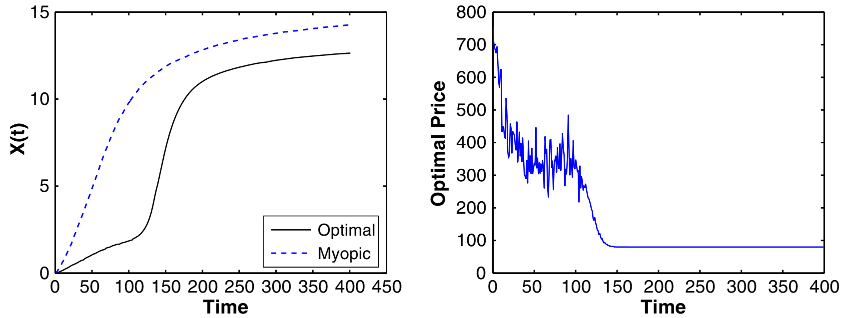
In this setting, we consider the case where both the uncertainty and the cost decrease with the sales level. The uncertainty varies as $\sigma^2(X_j) = \sigma_0^2(10 - X_j)$. The value of $\sigma_0$ was chosen to be 0.2. The form of $c(X_j)$ is the same as with the previous setting, that is, $c(X_j) = 100 - 0.2X_j$. The results are shown in Fig. 11. The optimal price trajectory gradually decreases with time. Table 3 shows that our algorithm gives better results than the myopic policy in this case as well.

**Remark 2.** It is important to note that our algorithm is a simulation-based method and can work even when information on the model is not known. Thus, for instance, our algorithm will continue




**(a)** State trajectory **(b)** Optimal price path

Fig. 9. $\sigma_0 = 0.5$.



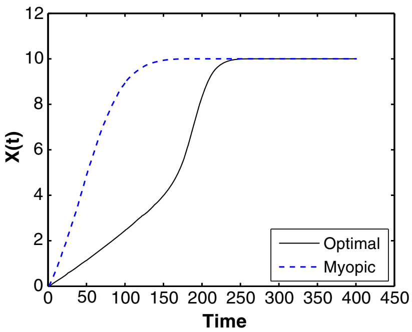

**(a)** State trajectory **(b)** Optimal price path

Fig. 10. $\sigma_0 = 0.8$.

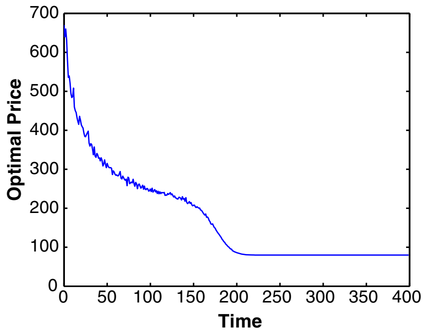
Table 2
Objective function values for the optimal and myopic prices with 95% confidence interval of mean values for the setting of cost learning with constant uncertainty

| $\sigma_0$ | Optimal price path | Myopic price |
|---|---|---|
| 0.1 | $-670.088 \pm 6.15832$ | $-472.557 \pm 0.9408$ |
| 0.3 | $-1146.342 \pm 26.29928$ | $-591.445 \pm 2.24028$ |
| 0.5 | $-1844.380 \pm 80.0268$ | $-696.259 \pm 3.26536$ |
| 0.8 | $-1141.415 \pm 24.527048$ | $-590.828 \pm 2.25988$ |

to work when no knowledge of the various parameters in the drift $b(\cdot, \cdot)$ as well as the diffusion $\sigma(\cdot)$ terms is available, as long as the state of the SDE $X_n$ is made known at each instant $n$ either via a simulation device or real data. This is however not the case with the myopic policy from Raman and Chatterjee (1995), where such information on the various parameters is critically required. Our algorithm will continue to work with any model other than the generalized Bass model as well. Thus, our technique is completely model free and also works on high-dimensional settings. In particular, note here that our setting requires optimizing a 400-dimensional parameter (the value of $N$). Our algorithm is able to perform fast computations because of the use of the efficient smoothed functional gradient estimator as a result of which one requires only two system simulations at each




**(a)** State trajectory




**(b)** Optimal price path

Fig. 11. Decreasing uncertainty and cost.

Table 3
Objective function values for the optimal and myopic prices with 95% confidence interval of mean for the setting of decreasing uncertainty with cost learning

| Optimal price path | Myopic price |
|---|---|
| $-1780.468 \pm 28.43568$ | $-409.886 \pm 0.049$ |

instant to update the entire parameter vector (of any dimension). In the case when one has access to real data, the variant of the algorithm briefly described in Remark 1 may be used.

## 7. Conclusions and future work

We studied the problem of computing the optimal price policy of a product over its life cycle in the market in the presence of demand dynamics and uncertainty. We proposed a highly efficient, simulation-based algorithm to compute the optimal pricing decisions at discrete time instants over the period of the planning horizon. We gave a proof of convergence of our algorithm. Our experiments illustrate that our algorithm is efficient even in high-dimensional settings involving parameters of dimension 400. Our algorithm is efficient because it incorporates the smoothed functional gradient estimator as a result of which, it requires only two system simulations to update the entire parameter at each instant. Our algorithm is also model-free and will work even when knowledge of system parameters is not available, as long as the state of the SDE is observed either through a simulation device or through real data. Also, our algorithm will continue to work with any model other than the generalized Bass model as well. This is unlike the myopic policy derived from Raman and Chatterjee (1995) that crucially depends on the model parameters. Numerical experiments showed that our algorithm gave better results than the aforementioned myopic policy.

As future work in this area, one can look at more complex models of product diffusion. For example, one can look at competitive scenarios where two rival firms introduce a similar product in the market and both have to advertise their products. Such situations have been modeled using the framework of "stochastic differential games" (Prasad and Sethi, 2004). Equilibrium strategies of stochastic differential games can be solved using extensions of simulation-based optimization methods. Another direction of future work could be looking at product diffusion over social networks. A typical social network will have sparsely connected vertices with nonuniform edge weights. One can also look at what the notions of optimal pricing mean in such models.

## References


Araman, V.F., Caldentey, R., 2009. Dynamic pricing for nonperishable products with demand learning. *Operations Research* 57, 5, 1169–1188.

Bass, F., 1969. A new product growth model for consumer durables. *Management Science* 15, 5, 215–227.

Bass, F., Krishnan, T., Jain, D., 1994. Why the Bass model fits without decision variables. *Marketing Science* 13, 3, 203–223.

Bhatnagar, S., 2007. Adaptive Newton-based smoothed functional algorithms for simulation optimization. *ACM Transactions on Modeling and Computer Simulation* 18, 1, 2:1–2:35.

Bhatnagar, S., Karmeshu, K., Mishra, V., 2009. Optimal parameter trajectory estimation in parameterized SDEs: an algorithmic procedure. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 19, 2, 8:1–8:27.

Bhatnagar, S., Prasad, H.L., Prashanth, L.A., 2013. *Stochastic Recursive Algorithms for Optimization: Simultaneous Perturbation Methods*. Springer, London.

Borkar, V.S., 1995. *Probability Theory: An Advanced Course*. Springer, New York.

Chinthalapati, V.R., Yadati, N., Karumanchi, R., 2006. Learning dynamic prices in multiseller electronic retail markets with price sensitive customers, stochastic demands, and inventory replenishments. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 36, 1, 92–106.

Farias, V.F., Van Roy, B., 2010. Dynamic pricing with a prior on market response. *Operations Research* 58, 1, 16–29.

Fourt, L.A., Woodlock, J.W., 1960. Early prediction of market success for new grocery products. *The Journal of Marketing* 25, 31–38.

Glasserman, P., 2003. *Monte Carlo Methods in Financial Engineering*, Vol. 53. Springer, New York.

Hirsch, M.W., 1989. Convergent activation dynamics in continuous time networks. *Neural Networks* 2, 331–349.

Jain, D., Mahajan, V., Muller, E., 1991. Innovation diffusion in the presence of supply restrictions. *Marketing Science* 10, 1, 83–90.

Jain, D.C., Rao, R.C., 1990. Effect of price on the demand for durables: modeling, estimation, and findings. *Journal of Business & Economic Statistics* 8, 2, 163–170.

Kalish, S., 1985. A new product adoption model with price, advertising, and uncertainty. *Management Science* 31, 12, 1569–1585.

Krishnan, T.V., Bass, F.M., Jain, D.C., 1999. Optimal pricing strategy for new products. *Management Science* 45, 12, 1650–1663.

Kushner, H.J., Clark, D.S., 1978. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Springer Verlag, New York.

Lin, K.Y., 2006. Dynamic pricing with real-time demand learning. *European Journal of Operational Research* 174, 1, 522–538.

Mahajan, V., Muller, E., Bass, F.M., 1990. New product diffusion models in marketing: a review and directions for research. *The Journal of Marketing* 54, 1–26.

Mahajan, V., Peterson, R.A., 1978. Innovation diffusion in a dynamic potential adopter population. *Management Science* 24, 15, 1589–1597.

Mansfield, E., 1961. Technical change and the rate of imitation. *Econometrica: Journal of the Econometric Society* 29, 4, 741–766.

Prasad, A., Sethi, S.P., 2004. Competitive advertising under uncertainty: a stochastic differential game approach. *Journal of Optimization Theory and Applications* 123, 1, 163–185.

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.

## Appendix: Myopic price policy

From Equation (6) in Raman and Chatterjee (1995), the optimal price trajectory can be written as

$$p_r(X) = c(X) - \frac{b(X, p_r)}{b_{p_r}(X, p_r)} - V_X, \tag{A1}$$

where, $b_{p_r}(X, p_r)$ is the derivative of $b(X, p_r)$ with respect to $p_r$ and $V_X$ is the derivative of the value function.

Substituting the expression for $b(X, p_r)$ in the above equation from the model in (8), we obtain

$$p_r(X) = c(X) - \frac{1 - \gamma p_r}{-\gamma p_r} - V_X, \quad \text{or,} \quad p_r(X) = \frac{1}{2}\left[c(X) + \frac{1}{\gamma} - V_X\right]. \tag{A2}$$

Now, the price trajectory for a myopic policy that treats the value function to be the same for all states will be

$$p_r(X) = \frac{1}{2}\left[c(X) + \frac{1}{\gamma}\right]. \tag{A3}$$