

Pricing Information Bundles in a Dynamic Environment

Jeffrey O. Kephart* Christopher H. Brooks† Rajarshi Das* Jeffrey K. MacKie-Mason‡
Robert Gazzale‡ Edmund H. Durfee†

Abstract

We explore a scenario in which a monopolist producer of information goods seeks to maximize its profits in a market where consumer demand shifts frequently and unpredictably. The producer is free to set an arbitrarily complex *price schedule*—a function that maps the set of purchased items to a price—but without direct knowledge of consumer demand it cannot compute the optimal schedule. Instead, it must employ a form of optimization based on trial and error. By means of a simple model of consumer demand and a modified version of a simple nonlinear optimization routine, we study a variety of parametrizations of the price schedule and quantify some of the relationships among learnability, complexity, and profitability. In particular, we show that fixed pricing or simple two-parameter dynamic pricing schedules are preferred when consumer demand shifts frequently, but that dynamic pricing based on more complex schedules tends to be most profitable when consumer demand shifts very infrequently.

1 Introduction

The relatively low cost of creating customized bundles of information goods opens up the potential for digital information providers to explore a richer and more complex set of pricing schemes than are traditionally applied to physical goods. For example, consider a personalized electronic journal. Rather than simply offering a single price per article, or a single price for a fixed

set of articles (such as an issue or a subscription), the publisher could offer a variety of arbitrarily complex volume discounts, or even more complex pricing schemes that take article categories into account.

While it seems natural that flexibly configurable sets of information goods should be priced in a comparably flexible manner, complex pricing schemes are only likely to come into widespread use if they are acceptable to purchasers. We anticipate that purchasers will receive assistance from software agents that will help analyze and respond to dynamic and complex pricing schemes.

Prior authors [4, 2, 13] have studied the benefit of various pricing schemes, based on the assumption that the producer knows the distribution of consumer preferences. Typically, however, producers do *not* have complete knowledge of all relevant parameters of this distribution. Instead, they are likely to refine an initial estimate of the mapping between prices and profits via trial and error. A trial price schedule yields, not just a profit, but also incremental information about the underlying consumer preferences, which could be valuable if the producer can capitalize on this information to attain higher profits in the future. Given the potential value of this information, a good pricing strategy will be a dynamic balance between *exploitation* of prices that are known from experience to be among the most profitable and *exploration* [12] of new prices in a quest for greater profitability.

Further complicating the challenge of developing a good dynamic pricing strategy is the fact that consumers' preferences may vary over time. This may stem from mass entry and exit of individual consumers, or from evolution of individual consumer tastes, and may be a particularly strong effect if those tastes are correlated, i.e. the consumers are susceptible to fads. Shifting (and, again, correlated) consumer preferences may also reflect an underlying change in the nature of the goods being offered by the producer, or in the way that they are combined with one another. Shifts in demand could also result from other changes in the market, such as the entry and exit of competitors offering similar or

*Institute for Advanced Commerce, IBM Research, PO Box 704, Yorktown Heights, NY 10598

†Artificial Intelligence Laboratory, University of Michigan, Ann Arbor, MI 48109

‡School of Information and Department of Economics, University of Michigan, Ann Arbor, MI 48109

partially substitutable goods.

Previously [3], we studied a scenario in which a producer learns more or less complex price schedules while facing a static population. From the learning transient, we inferred that the time scale on which preferences changed was likely to have an effect on the relative profitability of various price schedules.

In this paper, we focus directly on the question of how a monopolist might dynamically price information bundles in an environment in which demand shifts frequently and unpredictably. Following a brief account of related work in Section 2, we introduce a model that captures the essence of this problem in Section 3. We introduce several price schedules that range in complexity from 1 to 100 parameters, and describe a simple nonlinear optimization algorithm that exploits and explores the space of price parameters. In Section 4 we characterize the underlying profit landscapes pertaining to each type of schedule. We derive the optimal settings of the parameters, which could be obtained by a fully-informed producer; this serves as a useful baseline. Then, in Section 5, we present several simulations that establish relationships among the complexity, learnability, and profitability of the various price schedules, and identify the conditions under which particular price schedules are favored over others. In general, we find that, when consumer demand shifts frequently, fixed pricing based on complex schedules or dynamic pricing based on simpler schedules is preferred, but dynamic pricing based on complex schedules is advantageous when consumer demand shifts very infrequently. We conclude in Section 6 with a summary of our findings and some thoughts about challenges that lie ahead.

2 Related work

Most of the bundling literature assumes that the producer knows the consumer demand *a priori*. However, there have been some previous studies of how a firm can use price experimentation to learn consumer preferences. Grossman, Kihlstrom and Mirman [7] were among the first to quantify the extent to which a firm may experiment with prices that are not myopically optimal, thus giving up short-run profits, in order to learn the slope of the demand curve it faces. Subsequent authors have attempted to generalize the problem and identify the conditions under which “complete” learning will occur in the long run in stationary environments. For example, Aghion, Bolton, Harris and Jullien [1] determine such conditions, which include payoff functions that are smooth and quasi-concave, but note that neither adequate nor inadequate learning is generic.

The existing literature almost universally assumes that the preferences to be learned are stationary. In a

rare exception, Keller and Rady [8] identify the path of optimal actions for a monopolist facing an unknown demand curve that alternates between two states according to a continuous time Markov process. They find that for a given level of noise and a given discount rate, low rates of state changes will lead to “extreme” experimentation where the monopolist tracks the true state rather well, while higher rates of state changes will lead to “moderate” experimentation where the monopolist chooses actions near the action associated with the long-run average state and thus learns very little of the current true state of the world. They find that the transition from one experimentation regime to the other is discontinuous.

The machine learning community has addressed the problem of learning nonstationary functions. Early examples of this include reinforcement learning algorithms such as Q-learning and TD-learning [11], which can track a nonstationary function, and signal-processing algorithms such as Parzen windows [5], which can learn a changing probability distribution without the use of a parametric model. While these methods are able to successfully learn a nonstationary function in the limit, they require a large amount of data (a problem when data collection is costly), and neither explicitly addresses the exploration/exploitation tradeoff. Strategies for exploring optimally have also been examined in machine learning [12] and optimal control [6]. Typically these approaches either use a different definition of optimality (e.g., each data point is equally costly) or more restrictive assumptions about the nature of the problem to be optimized, such as stationarity.

3 Model

In this section we describe our model of a monopolist producer offering information goods to a population of C consumers. During each of a large number of discrete time steps, the producer has N new and unique information items, called articles, that it can offer to each consumer. At the beginning of each period, the producer announces a *price schedule* $P(n)$ (also denoted \vec{P}) that defines the price that any consumer must pay for n articles, where $0 \leq n \leq N$.¹ Each consumer c purchases a number of articles n_c that maximizes the surplus $S_c(n_c) = V_c(n_c) - P(n_c)$, where $V_c(n)$ represents c 's valuation for n articles. The producer then receives a profit π , which is a function of the price schedule \vec{P} and the consumers' valuations. More specifically, we assume (as is customary for information goods) that the marginal cost of duplicating, bundling, and deliver-

¹We assume that the producer charges the same price to all consumers because it is unable to charge customer-specific prices, either for legal or technical reasons (i.e., it may be unable to track individual consumers across transactions, or unable to capitalize on observations of individual behavior).

ing articles is zero, so the producer's marginal profit is simply $\pi(\vec{P}) = \sum_{c=1}^C P(n_c)$.

An experimental run consists of running this cycle for (typically) millions of time steps, after having chosen a particular set of conditions, such as the number of consumers C , a set of parameters describing their preferences, the type of price schedule, etc. The producer may experiment continually with $P(n)$ in an effort to optimize π . Continual experimentation is encouraged by periodic shocks in the consumer demand, which occur at an interval T . Typically, 10,000 to 100,000 shocks occur during a given run. For each run, we observe the sequence of \vec{P} and π . Of particular interest is the process of profit recovery that occurs after each shock, which we gauge by computing the average profit at time τ after a shock. In other words, we overlay and average the recovery periods to get an understanding of how successfully the producer can recuperate from a sudden shift in the aggregate consumer demand.

Having given this brief synopsis of the model, we now provide further details on the price schedule \vec{P} , the model of consumer preferences and their variation in time, and the process by which the producer attempts to optimize its profit in this dynamic environment.

3.1 Price schedules

In general, the price schedule $P(n)$ may be an arbitrary nonlinear relation between n and P . In practice, it is often convenient and efficient to parametrize $P(n)$ in terms of a small number of variables.² In this work we examine five such functional forms, ranging in complexity from 1 to N parameters:

- Pure Bundling. Consumers pay a fixed price b for access to all N articles.
- Linear Pricing. Consumers pay a fixed price p for each article purchased.
- Two Part Tariff. Consumers pay a subscription fee f , along with a fixed price p for each article.
- Mixed Bundling. Consumers have a choice between a per-article price p and a bundle price b .
- Nonlinear Pricing. Consumers pay a price P_i for i articles. An alternative representation is the incremental price vector $p_i \equiv P_i - P_{i-1}$, with P_0 defined to be zero.

3.2 Consumer preferences and demand shocks

We model consumers' preferences using a simple two-parameter model introduced by Chuang and Sirbu [4].

²Of course, $P(0)$ is always taken to be 0, to ensure that consumers are not charged unless they purchase at least one article.

In this model, consumer c 's valuation of its most preferred article is w_c , and the value of its i th most preferred article, v_{ci} , decreases linearly with i , as follows:

$$v_{ci} = \begin{cases} w_c(1 - \frac{i-1}{k_c N}) & \text{if } i - 1 \leq k_c N \\ 0 & \text{if } i - 1 > k_c N. \end{cases} \quad (1)$$

Thus consumer c places a positive value on $[k_c N] + 1$ articles³, unless $k_c > 1$, in which case the number of articles with positive valuation is capped at N . By summing the incremental valuations in Eq. 1, we obtain the overall valuation for n articles:

$$V_c(n) = \sum_{i=1}^n v_{ci} = w_c n \left(1 - \frac{n-1}{2k_c N}\right). \quad (2)$$

As would be expected, $V_c(0)$ is zero, signifying that consumers ascribe zero value to zero articles. The number of articles n_c purchased by consumer c therefore depends simply on the price schedule $P(n)$ and c 's valuation function $V_c(n)$. Specifically, n_c is the number of articles that maximizes $V_c(n) - P(n)$.

For purposes of analytic tractability, we assume that w_c is fixed at a universal value w for all consumers, but we allow k_c to be both heterogeneous and time-varying. Specifically, w is fixed at 10, and k_c is drawn uniformly from between 0 and $\bar{k} = 0.7$. Every T time steps, all of the consumers' valuations are redrawn from this distribution, resulting in the aforementioned shock in the aggregate consumer demand.

3.3 Dynamic optimization of the price schedule

If the producer knew the consumer demand at all times, then it could simply respond to a demand shock by immediately recomputing and announcing the new optimal price schedule $P(n)$. However, it is unrealistic to assume that the producer is so well informed. In this paper, we are primarily interested in how a producer can re-learn the mapping between price and profit, and use this as a basis for resetting $P(n)$ to yield more (if not optimal) profit.

We model the producer's ignorance of aggregate consumer demand by assuming that the producer does not know the distributions of w_c and k_c ; in fact, it does not even know about the Chuang-Sirbu structure of the consumer demand. Instead, the producer is limited to observing its own history of price schedules \vec{P} and corresponding profits $\pi(\vec{P})$, and from this it tries to learn as profitable a price schedule as possible. Inherent in this task is a tradeoff between exploration and exploitation: the producer needs to explore the function $\pi(\vec{P})$ in an effort to find better schedules than it has yet found, but it also wishes to capitalize on schedules that are already known to be among the most profitable.

³Here $[x]$ denotes the integer just less than or equal to x .

To model the producer’s process of relearning and reoptimizing, we employ a modification of the *amoeba* algorithm, a variant of the *simplex algorithm* [10] for unconstrained nonlinear optimization problems.⁴ At each time step, the amoeba algorithm maintains a nondegenerate simplex, an n -dimensional convex hull of $n+1$ vertices, $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_n$, and their respective function values. In our application, each vector \vec{x} represents a setting of the price parameters appropriate for the given price schedule, observed at an earlier time, and it is associated with a profit $\pi(\vec{x})$ that was obtained for that setting of the price parameters. Amoeba alternately reflects, contracts, and expands the simplex in an attempt to locate an optimum in a function’s landscape. At each time step, amoeba chooses a new point \vec{x} , which is translated into $P(n)$, and the corresponding profit $\pi(\vec{x})$ is determined.⁵

An attractive feature of amoeba for our application is that, unlike many other nonlinear optimization algorithms, it does not require any gradient information. Without access to the underlying consumer demand, or even a knowledge of its functional form, a producer can only sample the profit landscape directly.

An unattractive feature of amoeba is its implicit assumption that the landscape is static. This can cause it to fail spectacularly when the landscape is dynamic [9]. The problem begins with amoeba’s failure to detect changes in the landscape $f(\vec{x})$. For example, suppose that, in some neighborhood including the current best point \vec{x}_0 , with functional value $f(\vec{x}_0) = z_0$, there is a sudden general depression in f such that all points are shifted down by a constant c . Then amoeba will see this depression for newly sampled points near x_0 , but will steadfastly continue to believe that $f(\vec{x}_0) = z_0$ rather than $z_0 - c$. Amoeba will continue to search in progressively smaller neighborhoods around x_0 , and will reach the conclusion that there is an arbitrarily sharp discontinuity there. Another problem is that convergence of the simplex around the optimum, while useful in static environments, makes amoeba extremely unresponsive in dynamic situations because it takes too long for the simplex to re-expand to a size that permits efficient exploration of the changed landscape.

We circumvent the problem of detecting changes in the landscape by directly resetting the simplex whenever such changes occur. While this is admittedly artificial, essentially the same effect could be achieved more naturally by modifying the amoeba algorithm to include occasional resampling. To counteract overconvergence of the simplex, we reset it as follows whenever the landscape changes. For each component x_{ij} of the vertex \vec{x}_i , choose a new value x'_{ij} at random from the uniform distribution $U[(1 - \phi)x_{ij}, (1 + \phi)x_{ij}]$. Thus, the parameter

⁴This simplex algorithm should not be confused with the simplex algorithm for linear programming.

⁵For a detailed survey on amoeba refer to Wright [14].

ϕ governs the amount of exploration that takes place immediately after a demand shock. Appropriate values for ϕ must be determined by experiment. If ϕ is too small, the simplex is more susceptible to entrapment in local optima that are significantly lower than the global optimum. On the other hand, if ϕ is too large, amoeba will have to start from scratch, and may not converge as quickly as it would have had it started from a point relatively close to the new global optimum.

We have made one additional modification to avoid the occasional problem of ending up in a zero-profit plateau. This happens when all of the simplex vertices have price parameters that are set so high that no consumers wish to purchase any articles. To prevent this from occurring, we divide all price parameters by two whenever this condition is detected, repeating the procedure until a positive profit is obtained.

4 Price schedules and profit landscapes

In this section, we compute the optimal profit for each of the five price schedules introduced in Section 3. For population sizes $C = 1$ and $C \rightarrow \infty$, we derive a good analytic approximation; for finite $C > 1$ we compute the optimal profits numerically via simulation.

Next, to gain insight into the learning and optimization task faced by a producer who lacks *a priori* knowledge of consumer demand, we illustrate typical *profit landscapes*. (A profit landscape is the functional dependence of the profit upon the price parameters of a given type of price schedule.) The computed optimal profits are the global optima of these landscapes, and therefore establish an upper limit on the profit that can be attained by a producer, regardless of how well-informed it is about consumer demand.

4.1 Optimal profits

4.1.1 $C = 1$ (perfectly correlated consumers)

First, consider the case $C = 1$. It is useful to think of $C = 1$ as representing, not just a single consumer, but a cluster of many consumers that share exactly the same value of k . When a shock occurs, a new value of k is chosen and adopted by all of the consumers in unison. Analytic expressions for the resulting optimal price parameters \vec{P}^* (as a function of k) and the optimal profits π^* (averaged over the uniform distribution $U[0, \bar{k}]$) are summarized in the first two columns of Table 1.

We find that, for all of the price schedules except linear pricing, the producer can set its price parameters to extract all of the social surplus. The maximal social surplus is obtained when the consumer receives all articles for which she has a positive valuation. According to Eq. 1, the number of such articles is $n = \lfloor kN \rfloor + 1 \approx kN + 1$, to within a factor of $O(1/N)$.

Schedule	$C = 1$		$C \rightarrow \infty$	
	\vec{P}^*	π^*	\vec{P}^*	π^*
Pure Bundle	$b = \frac{wkN}{2}$	$\frac{w\bar{k}}{4}$	$b = \frac{w\bar{k}}{4}$	$\frac{w\bar{k}}{8}$
Linear	$p = \frac{w}{2}$	$\frac{w\bar{k}}{8}$	$p = \frac{w}{2}$	$\frac{w\bar{k}}{8}$
Two-Part	$f = \frac{wkN}{2}, p = 0$	$\frac{w\bar{k}}{4}$	$f = \frac{2w\bar{k}N}{27}, p = \frac{w}{3}$	$\frac{4w\bar{k}}{27}$
Mixed	$b = \frac{wkN}{2}, p \rightarrow \infty$	$\frac{w\bar{k}}{4}$	$b = \frac{8w\bar{k}N}{27}, p = \frac{2w}{3}$	$\frac{4w\bar{k}}{27}$
Nonlinear	$p_i = \frac{wkN}{2} \delta_{i1}$	$\frac{w\bar{k}}{4}$	p_i given by Eq. 3	$\frac{w\bar{k}^2 ((1+\bar{k}) \ln(1/\bar{k}) - 2(1-\bar{k}))}{(1-\bar{k})^3}$
Perfect	—	$\frac{w\bar{k}}{4}$	—	$\frac{w\bar{k}}{4}$

Table 1: Optimal price parameters and profits per article per consumer for 5 selected price schedules; $C = 1$ and $C \rightarrow \infty$. The profit attainable via perfect quantity-based price discrimination is shown for comparison. For $C = 1$, the optimal price parameters \vec{P}^* are expressed as a function of the k -value of the consumer, while the optimal profits π^* pertain to an average over $U[0, \bar{k}]$. For $C \rightarrow \infty$, optimal prices and profits are based on averages over $U[0, \bar{k}]$. Formulas are valid as $N \rightarrow \infty$; finite N corrections are of $O(1/N)$. In the expression for the optimal nonlinear price parameters for $C = 1$, δ_{i1} represents the Kronecker delta function, equal to 1 if $i = 1$ and 0 otherwise.

From Eq. 2, the value of these articles is $\frac{wkN}{2}$, to within a factor of $O(1/N)$.

By setting a bundle price of $b = \frac{wkN}{2}$, the producer can extract all of this surplus⁶, and therefore its profit per consumer⁷ will also be $\pi = \frac{wkN}{2}$. Since all price schedules except linear pricing can express pure bundling as a degenerate case, they can all implement this optimal solution, as summarized in Table 1. Only half as much surplus can be achieved by linear pricing.

In our model of periodic shock and recovery, a perfectly informed producer employing any of our set of price schedules other than linear pricing could use the above analysis to compute the optimal parameters, obtaining a profit of $\frac{wkN}{2}$ for the duration of each recovery period. Given that the value of k shared by all consumers is drawn from a uniform distribution ranging from 0 to \bar{k} , the expected time-averaged profit would be $\frac{w\bar{k}N}{4}$. In a time-averaged sense, such a producer would be practicing perfect quantity-based price discrimination, always charging just as much as each consumer would bear. For linear pricing, the time-average profit would be exactly half this amount.

4.1.2 $C \rightarrow \infty$ (perfectly uncorrelated consumers)

Now suppose that each consumer draws k independently of all others, and that the number of consumers $C \rightarrow \infty$. Then the optimal price parameters and corresponding

⁶This makes the usual assumption that an indifferent consumer will purchase the bundle, which conveniently avoids infinitesimal quantities.

⁷Henceforth, the *profit per consumer* will be simply referred to as the *profit*.

profits per consumer can be computed to an approximation accuracy of $O(1/N)$. Table 1 summarizes the results first derived by Brooks et al. [3]. All results are normalized by the bundle size N . The formula for the optimal nonlinear price schedule given by Brooks et al. [3] is incorrect; the correct form is given by Equation 3, in the appendix.

4.1.3 Optimal profits for finite N and C

The optimal profits for each price schedule for a range of values of N and C are tabulated numerically in Table 2. Each consumer is assumed to have Chuang-Sirbu parameters $w = 10$ and k drawn from $U[0, 0.7]$. The values for infinite bundle size and $C = 1$ and $C \rightarrow \infty$ are computed using the analytic expressions in Table 1. The values for finite N and C are computed via simulation, as follows. Each of 10,000 landscapes is generated by drawing C k -values from $U[0, 0.7]$. An excellent approximation to the absolute peak for each landscape is obtained by running the amoeba nonlinear optimization algorithm to 10-digit convergence from 100 randomly chosen positions on the landscape (clustered loosely around the value obtained from Table 1) and identifying the highest peak found over the 100 runs. (Experimentation establishes that this is more than sufficient to find the global peak for the two-part tariff landscape.) The peak profits are then averaged over the 10,000 landscapes. Finally, the values for $N = 100$ and $C \rightarrow \infty$ are computed using two different techniques. For two-part tariff, we developed an exact analytic expression for the $C \rightarrow \infty$ landscape as a function

Schedule	$N \rightarrow \infty$ $C = 1$	$N = 100$ $C = 1$	$N = 100$ $C = 10$	$N = 100$ $C = 100$	$N = 100$ $C \rightarrow \infty$	$N \rightarrow \infty$ $C \rightarrow \infty$
Pure Bundle	1.7500 (1.0104)	1.7946 (1.0066)	1.1220 (0.2839)	0.9497 (0.0878)	0.9010	0.8750
Linear	0.8750 (0.5052)	0.9208 (0.5032)	0.9103 (0.1608)	0.9044 (0.0515)	0.8999	0.8750
Two-Part	1.7500 (1.0104)	1.8029 (1.0115)	1.2322 (0.2732)	1.1043 (0.0777)	1.0710	1.0370
Mixed	1.7500 (1.0104)	1.8017 (1.0078)	1.2402 (0.2647)	1.1071 (0.0766)	1.0706	1.0370
Nonlinear	1.7500 (1.0104)	1.8129 (1.0093)	1.3934 (0.2716)	1.2564 (0.0842)	1.2049	1.1519

Table 2: Optimal profits for each price schedule for various values of bundle size N and consumer population C . Values are computed assuming that each consumer has Chuang-Sirbu parameters $w = 10$ and k drawn randomly from the uniform distribution $U[0, 0.7]$. In cases where C is finite, numbers in parentheses represent the standard deviation across different realizations of the profit landscape, as measured across 10,000 landscapes. The last column represents the theoretical optimal profits for $N \rightarrow \infty, C = 1$. The standard deviation across different realizations can be computed analytically; it is $\frac{wk}{4\sqrt{3}}$ for all schedules except linear pricing, where it is half of this value.

of f and p and ran amoeba on it 100 times to 10-digit convergence. Roughly 30% of the runs yielded the highest peak (to 10-digit accuracy), and this was identified as the optimal value. For other landscapes, we generated 100 instances of a $C = 10^6$ landscape, running amoeba multiple times to identify the global optimum. The variation across landscapes was no more than 0.1% to 0.2%, and these results were averaged over the 100 landscapes.

Several trends are apparent:

- For all price schedules, the finite size corrections of $O(1/N)$ always increase the profits above their analytic values. Experiments have confirmed that this trend continues for smaller values of N ranging down to $N = 10$.
- For all price schedules except linear pricing, the profits vary inversely with C . This is consistent with the analytic results for $C = 1$ and $C \rightarrow \infty$, shown in Columns 7 and 2, respectively. For linear pricing, the profits are invariant with respect to C . Intuitively, as C is reduced towards 1, it becomes easier for the producer to practice quantity-based price discrimination, as the consumers are more uniform in their valuations. Linear pricing fails to follow this trend because it does not support quantity-based price discrimination.
- The profit tends to increase with the number of parameters in the price schedule. Unless $C = 1$, however, no price schedule can perform as well as perfect quantity-based price discrimination.
- For large C , the two one-parameter schedules (pure bundling and linear) and the two two-parameter schedules (two-part tariff and mixed bundling) yield equal profits. As C is reduced, pure bundling emerges as superior to linear pricing, as has been noted. However, two-part tariff and mixed bundling

remain essentially equal in profitability throughout the entire range of C .

4.2 Profit landscapes

The results in Table 2 are obtained by finding the peak on a landscape defined by the price parameters. Since the producer does not know the consumers' valuations, it must explore this landscape in an effort to find the peak. The degree to which the producer is successful in finding the global optimum, or a close approximation to it, has much to do with the shape of the landscape.

Figure 1 shows representative landscapes for the two-part tariff and mixed bundling landscapes for $C = 1$, $C = 10$, and $C = 100$. The $C = 1$ landscapes consist of gradual slopes leading up to an infinitely sharp precipice. This reflects the ability of the producer to practice perfect quantity-based price discrimination for two-part tariff and mixed bundling. The optimal pure bundle price of $b \approx 175$ is clearly evident for $C = 1$ for both landscapes. The optimal profit is obtained when the price is just equal to the consumer's maximal willingness-to-pay; if the price is increased beyond this, the profit will be zero. The mixed bundling landscapes provide insight into the pure bundling and linear pricing landscapes, which are one-dimensional. In the limit of high per-article price p , mixed bundling is equivalent to pure bundling, and therefore the pure bundling landscape can be traced horizontally along the upper boundaries of the mixed bundling landscapes. In the limit of high pure bundle price b , mixed bundling reduces to linear pricing, and therefore this one-dimensional landscape can be traced vertically along the righthand boundaries of the mixed bundling landscapes. For $C = 1$, linear pricing is the only form that lacks a sharp cliff, and correspondingly it is the only one that does not support perfect quantity-based price discrimination.

For all of the price schedules, the landscapes for $C > 1$ can be understood as a superposition of the

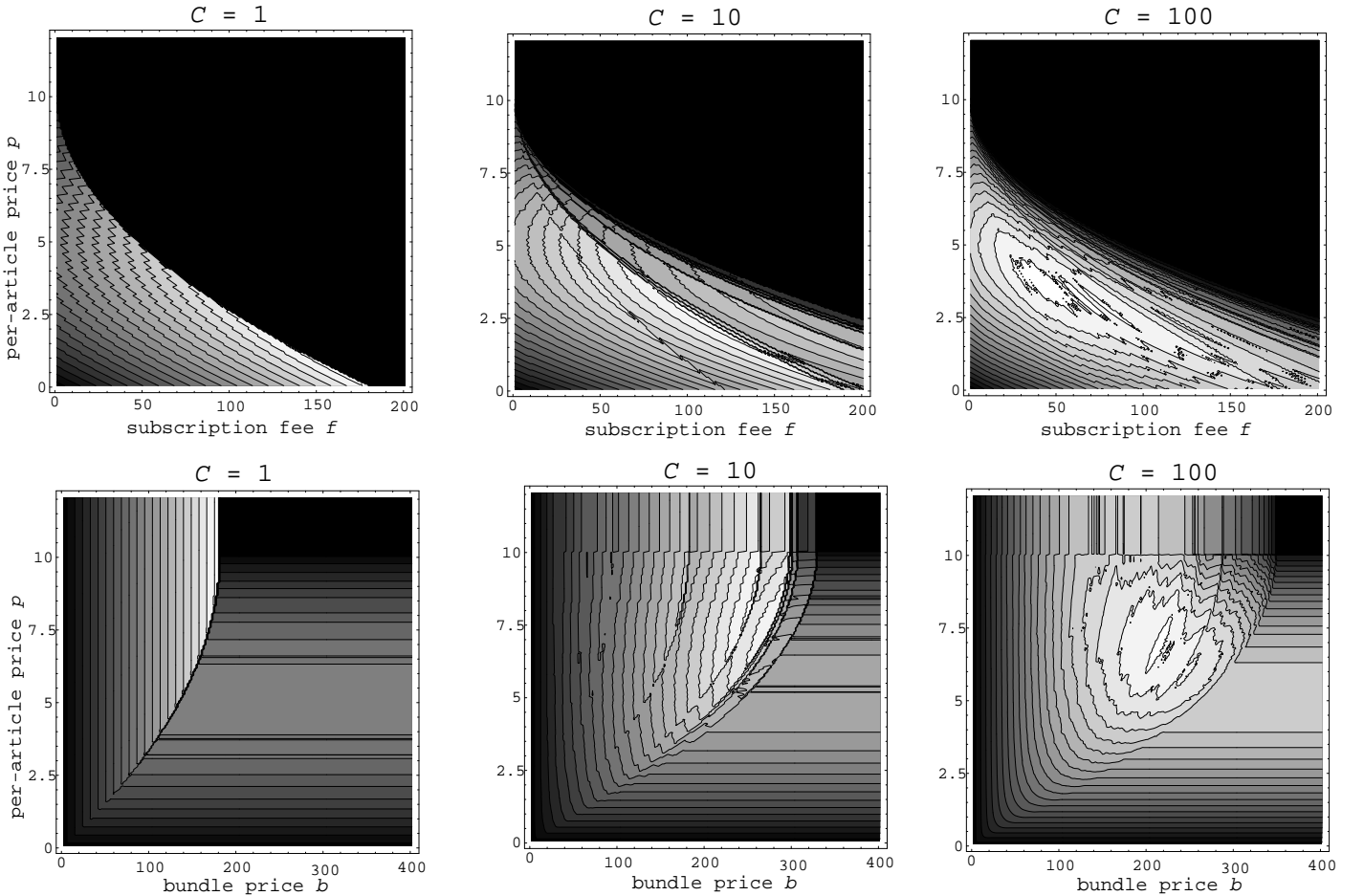


Figure 1: Contour plots of profit landscapes for two-part tariff and mixed bundling, with $N = 100$ and various consumer population sizes C . Lighter coloration indicates higher profits; scales are specific to each plot and therefore cannot be compared across different plots. Top row, left to right: Two-part tariff landscapes with $C = 1$, $C = 10$, and $C = 100$. Bottom row, left to right: Mixed bundling landscapes with $C = 1$, $C = 10$, and $C = 100$. For $C = 1$ landscapes, $k = 0.35$ (the average of the statistical distribution). For $C = 10$ and $C = 100$ landscapes, the k values are drawn at random from the distribution $U[0, 0.7]$.

individual $C = 1$ landscapes generated by C random draws from the distribution $U[0, 0.7]$. The resulting landscapes are peppered with myriad bumps and discontinuities. As C increases, the bumps and discontinuities increase in number but decrease in magnitude, eventually resulting in an arbitrarily smooth surface as $C \rightarrow \infty$. For $C = 100$, the bumps are small, but can still be hazardous to nonlinear optimizers like amoeba. Surprisingly, even for $C = 100$, the landscapes can shift substantially from one realization to another. The variability and bumpiness of the landscapes is illustrated in Fig. 2, which shows cross sections of two successive realizations of two-part tariff landscapes with $C = 100$. The peak position shifts substantially, from $(f, p) = (66.658, 3.145)$ to $(48.564, 3.497)$. On each landscape, the position of the optimal peak would be highly non-optimal on the other landscape.

5 Experimental results

5.1 Two-part tariff experiments

In our first series of experiments, we focus on the two-part landscape. For a fixed re-expansion radius ϕ and consumer population C , we set the initial price parameters to the theoretical values for $C \rightarrow \infty$, as given in Table 1, and then choose the initial simplex vertices at random from within the re-expansion radius. Although these formulas become progressively less accurate as N and C are reduced, they establish a reasonable starting point, and moreover all memory of the initial conditions is typically erased after a relatively small number of shocks. In a real price-setting scenario, a producer would presumably have a reasonable estimate of the right parameters, perhaps within a factor of two or so.

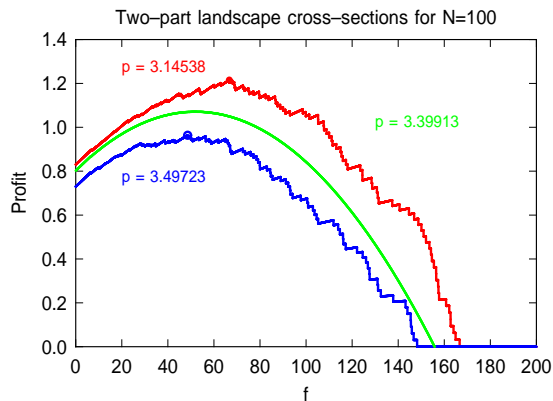


Figure 2: Cross sections of landscapes for $N = 100$. Middle curve represents the ideal $C \rightarrow \infty$ landscape; curves above and below it represent two realized landscapes for $C = 100$. Cross sections are for values of per-article parameter p fixed at the values indicated in the figure, which are chosen to be those at which the absolute peak in the landscape occurs.

Then, we ran a simulation as described in Section 3, in which the producer announces a new price parameter vector (f, p) at each time step, and uses the modified amoeba algorithm to set (f, p) at the next time step. Shocks were introduced periodically every $T = 10,000$ time steps. This allowed amoeba plenty of time to perform its optimization. During the experimental run, 100,000 shocks occurred, i.e. the run continued for 10^9 time steps. We measured the average profit as a function of time since the most recent shock by overlaying the recovery periods and averaging over them to produce a single characteristic profit transient.

Results for various values of ϕ and C are depicted in Figure 3. For $C = 1$, we find that, for all values of ϕ , the initial effect of a shock is quite severe—on average, the profit drops to about $1/3$ of the global optimum. This can be attributed to the great disparity between different realizations of the landscape for $C = 1$. (Recall that, according to the analysis, the peak positions are linear in k , which is drawn from $U[0, 0.7]$.) However, amoeba tends to find the new global optimum after a few hundred time steps. This is not surprising, as the $C = 1$ landscapes are relatively easy to negotiate, containing a single well-defined peak. What is surprising is that the characteristic profit transient, which is an average over 100,000 shocks, contains significant oscillations. One might expect a given individual profit recovery to exhibit oscillations as the amoeba flexes its pseudopods and gropes its way toward a new optimum, but it is surprising that phase cancellation does not wash out the oscillations when the average over shocks is performed.

Significantly, the path that amoeba takes to the new

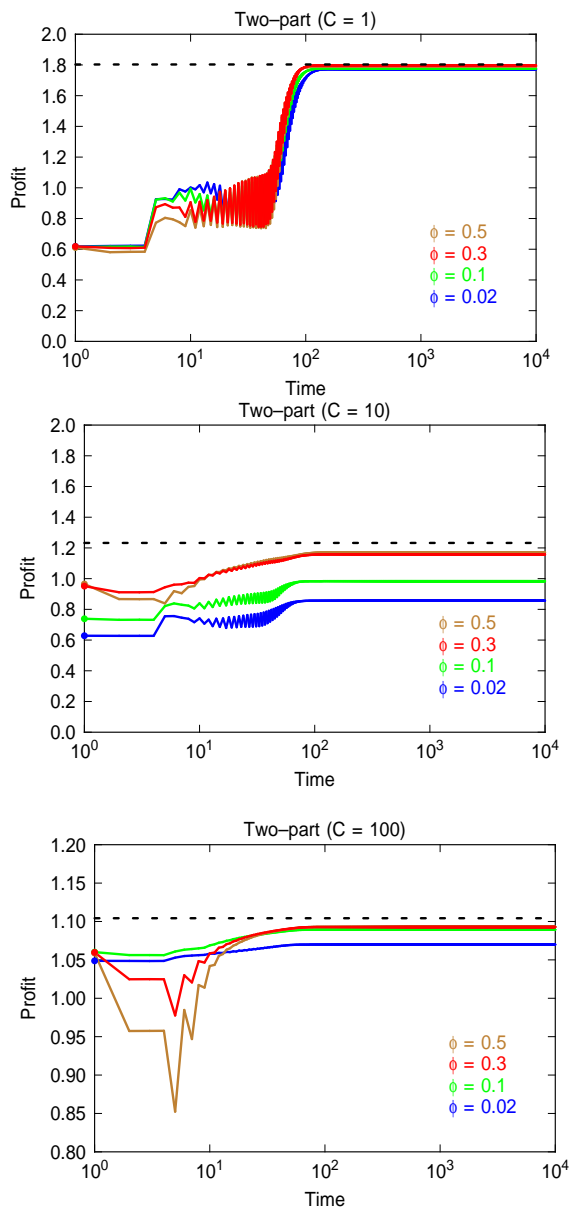


Figure 3: Learning transient for two-part tariff for various values of amoeba re-expansion parameter ϕ for a) $C = 1$, b) $C = 10$, and c) $C = 100$ (note expanded profit scale). Dashed horizontal line indicates optimal profit as given by Table 2. Note the logarithmic time scale.

peak is somewhat dependent on ϕ . Very soon after the shock, within 10 to 25 time steps, the smallest values of ϕ provide the quickest recovery in profit. However, from 25 to 200 time steps after the shock, the larger values of ϕ provide a somewhat higher profit. This suggests that $\phi \approx 0.02$ would be preferred if T were very short (in the range of 10 or so), while $\phi \approx 0.5$ would be most preferred if T were a few hundred or more.

The situation changes somewhat for $C = 10$. In this case, the landscape is a good deal more complex and craggy, with numerous substantial discontinuities. Even for the largest values of ϕ , and for very long times after the shock, amoeba attains an average profit of 1.171, as compared with the theoretical optimum value of 1.232 in Table 2, or about 95% of the optimum on average. Performance is quite poor for $\phi = 0.02$. Asymptotically, amoeba only reaches 0.857, or less than 70% of the global optimum, and the initial profit following a shock is only 0.628, compared with 0.965 for $\phi > 0.25$. Early in the simulation run, just after being launched from a position near the average position of the global optimum, the performance for small ϕ is quite reasonable, but with successive shocks the performance degrades as (f, p) diffuses away from this position towards one in which the per-article price p is very small, thus approximating pure bundling. One can infer that, regardless of T , there would never be a reason to use $\phi < 0.3$. For T less than a few hundred time steps, $\phi = 0.3$ is the preferred setting, but the slightly better asymptotic profit for $\phi = 0.5$ favors it when T exceeds a few hundred.

For $C = 100$, the larger values of ϕ again provide the best asymptotic profit, and are therefore favored for large T . The initial recovery for large ϕ is poorer than that for small ϕ because of the larger exploratory movements of the vertices of the simplex. Thus, for smaller values of T , $\phi \approx 0.1$ appears to be a good compromise, offering relatively good transient and asymptotic performance. This intermediate value of ϕ , and smaller values of ϕ as well, perform better for $C = 100$ than they do for $C = 10$. One reason is that the landscape is smoother, making it easier for amoeba to negotiate the landscape. This is reflected in the fact that smaller values of ϕ are more successful for $C = 100$ than for $C = 10$. A second reason is that individual realizations of the landscapes, and therefore the positions of their peaks, are closer to one another, reducing the need for extensive exploration. By confining the initial re-expansion to a smaller region around the old peak, amoeba can expect to extract higher profits in the short term.

A general trend as C increases is that the gains from re-learning are reduced. This is a simple consequence of the law of large numbers: draws of C different k values from a distribution become increasingly like one another as C grows, and therefore so do the landscapes based upon them. To see an appreciable learning effect

in such cases, the statistical distribution itself would need to drift—an interesting topic for future study.

5.2 Other price schedules

Figure 4 shows the profit recovery curves for the other four price schedules for the intermediate case $C = 10$, which is chosen because the landscapes are the most challenging.

For pure bundling, the dependence of the profit recovery curve on ϕ is similar to that for two-part tariff. Small values of ϕ are universally inferior up to about $\phi = 0.3$. While $\phi = 0.3$ offers a quicker recovery, $\phi \geq 0.5$ is preferred a few dozen time steps after the shock. Mixed bundling behaves similarly. Small values of ϕ are particularly poor for nonlinear pricing, leading to asymptotic profits of only about 1/3 of the optimal value, compared with about 94% for $\phi \geq 0.5$. For nonlinear pricing, the first 101 steps of the amoeba are simply function evaluations of the newly chosen simplex vertices; these show up as a trough in the profit recovery curve. Linear pricing, as usual, is anomalous. Regardless of ϕ , amoeba always comes very close to the global optimum, and does so very quickly, within about 10 time steps after the shock. The reason is that the linear pricing landscape is much easier to negotiate than the others, as it consists of a simple broad peak with small bumps. However, the quick learnability of linear pricing is not enough to make up for the fact that its global optimum is significantly lower than that of the other schedules (except pure bundling when $C \rightarrow \infty$; see Table 2). Even at times very soon after the shock, the other schedules offer greater profits.

5.3 Comparison of price schedules

Analysis of Fig. 4 in the previous subsection suggested that the best choice of ϕ depended on whether one was interested in short-term or long-term recovery. To gain further insight into this, we have run experiments with smaller values of T , and measured the time-averaged profit (the profit recovery curve averaged from times 0 to T after the shock).⁸ In each case, we have chosen the value of ϕ that maximizes this quantity.⁹ The results are summarized in Table 3.

As was observed previously, larger values of ϕ are favored when the interval between shocks is greater. In a fairly broad range, from extremely frequent shocks ($T = 10$) to fairly infrequent shocks ($T = 1000$), mixed bundling offers the highest average profit, as it can

⁸In general, a more appropriate figure of merit would be cumulative future discounted profit with a chosen discount rate; we are effectively assuming a discount rate of zero.

⁹The profits are only mildly sensitive to the exact value of ϕ ; for example, in the entry for Pure Bundle and $T = 10$, reducing ϕ from 0.3 to 0.2 reduces π by only 0.1%, while increasing ϕ to 0.4 reduces π by only 1.2%.

Schedule	$T = 10$	$T = 100$	$T = 1000$	$T = 10000$	Optimal	Opt. (fixed)
Pure Bundle	0.847 (0.30)	0.955 (0.90)	1.066 (0.95)	1.079 (1.00)	1.122	0.901
Linear	0.900 (0.02)	0.904 (0.02)	0.905 (0.10)	0.907 (0.10)	0.910	0.900
Two-Part	0.853 (0.20)	1.034 (0.50)	1.144 (0.60)	1.186 (0.80)	1.232	1.071
Mixed	1.016 (0.10)	1.090 (0.35)	1.146 (0.40)	1.154 (0.40)	1.240	1.071
Nonlinear	—	—	1.112 (0.50)	1.312 (0.50)	1.393	1.205

Table 3: Time-averaged profits for each schedule for several values of shock interval T , with $C = 10$. Optimal ϕ is in parentheses. Second-to-last column gives the optimal profits for $N = 100, C = 10$ from Table 2; this represents the upper limit on what can be obtained by a responsive, fully-informed producer. Last column gives the optimal profits for $N = 100, C \rightarrow \infty$ from Table 2; this represents an upper limit on what could be obtained by a non-responsive, fully-informed producer using a fixed price schedule.

be learned relatively quickly, yet yields good long-term profits. For extremely infrequent shocks (T is at least a few thousand time steps), nonlinear pricing offers a higher average profit than the other price schedules. This is attributable to the intrinsically higher global optimum for nonlinear pricing, combined with the ability of amoeba (properly tuned) to attain a profit of about 94% of that optimum in a dynamic setting, provided that plenty of time is given to reach it.

While it is clear from the profit recovery curves in Figs. 3 and 4 and from the trend towards larger profits for larger shock intervals in Table 3 that re-learning is both possible and valuable, it is of interest to compare re-learning and re-optimization with an alternative strategy. Suppose that the producer were to use the amoeba algorithm as described for a long time, recording the series of price parameters and profits and averaging them to obtain an averaged profit landscape. Then, after it was satisfied that it had a sufficient understanding of the average landscape, it could find the optimal price parameters for this surface. Afterwards, it would ride out any subsequent shocks with its price schedule fixed at the optimum for the time-averaged landscape. How would this strategy compare?

Experimentation has shown that, to within a few tenths of a percent, the values in Table 2 for $N = 100, C \rightarrow \infty$ provide an excellent approximation to the average profit that could be obtained with a fixed price schedule. This is reasonable, because the landscape for $C \rightarrow \infty$ is equivalent to an average over an infinite ensemble of landscapes with arbitrary finite C . Comparing the profits in the last column of Table 3 with the results for different shock intervals, we find that non-responsive pricing is a reasonable alternative to responsive pricing for sufficiently frequent shocks. Specifically, non-responsive pricing is best for $T < 20$ for pure bundling, $T < 500$ for two-part tariff, $T < 50$ for mixed bundling, and $T < 3000$ for nonlinear pricing. The progression from smaller to larger values of the threshold in T with increasing complexity of the price schedule reflects the increased difficulty of learning price sched-

ules with more parameters. All of these thresholds are very approximate, based on rough interpolation of the values in Table 3.

Two caveats are worth mentioning. First, the analysis of this section applies for what is effectively a discount rate of zero; a non-zero discount rate would place a greater emphasis on quick learnability. Second, while the values in the last column of Table 3 may seem to imply that fixed pricing can be very attractive under the right conditions, it should be recognized that fixed pricing can fail miserably if the assumption of stationarity in the distribution of consumer preferences is violated. The dynamic pricing schemes are not nearly so vulnerable to such an assumption, and are likely to do a much better job of tracking nonstationary consumer preferences.

6 Conclusion

In this paper, we have explored how a monopolist producer of information goods can optimize its price schedule without knowing the consumer demand *a priori*. Using a modified version of amoeba, a simple nonlinear optimization technique, the producer can track a shifting profit landscape with a moderate degree of responsiveness. We have characterized the producer's response to shocks in the aggregate demand, and related it to the topography of the profit landscape, the degree of exploration, and the frequency of the shocks. When demand shocks are very infrequent, the nonlinear pricing scheme is favored, and the producer can reach profits that come within about 95% of the profit that it could attain with *a priori* knowledge of the demand. As the frequency of shocks is increased, the profitability of any given schedule decreases, and less complex price schedules such as two-part tariff or mixed bundling are favored because they can be learned much more quickly. We have also established that, if shocks are sufficiently frequent, fixed pricing based on a learned statistical average of the landscape can be more profitable than continual re-learning of individual landscapes. Another

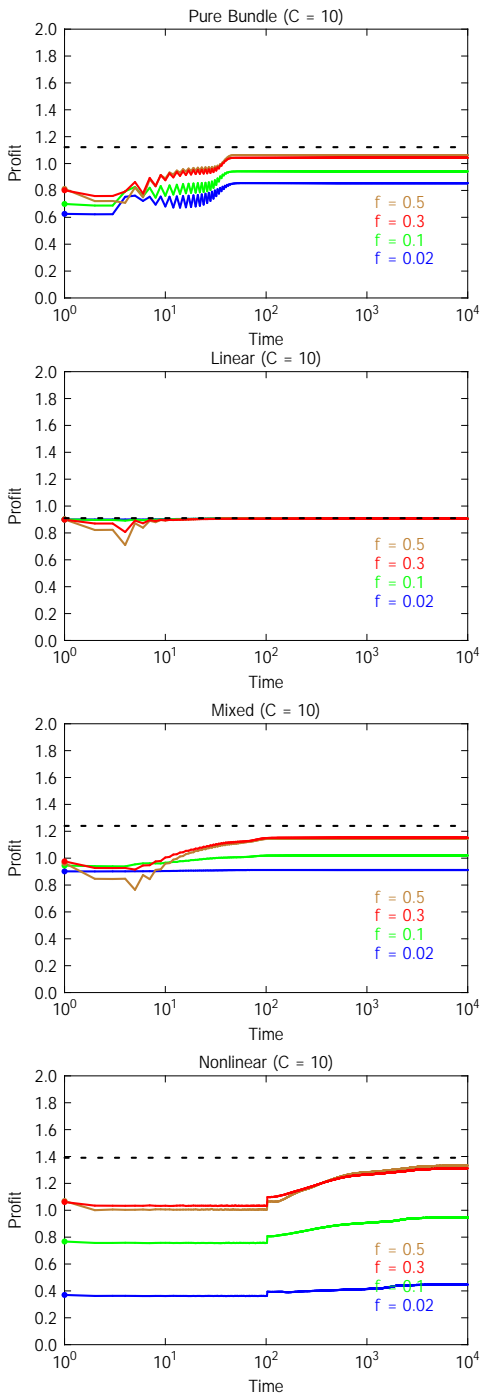


Figure 4: Learning transient for a) pure bundling, b) linear, c) mixed bundling, and d) nonlinear pricing for $C = 10$. Dashed horizontal line indicates optimal profit as given by Table 2. Note the logarithmic time scale.

finding is that, following the shock, a smaller degree of initial exploration by amoeba is desired as the frequency of shocks increases. This helps the producer capitalize on previous learning, although it can increase the probability of getting stuck at a less optimal local optimum.

The modified amoeba presented here is certainly not the best conceivable algorithm for dynamically pricing information bundles. Amoeba is attractive in that it requires only function evaluations, but despite our modifications it is ill-suited in several other ways. A better algorithm would detect and handle changing landscapes more gracefully, and would offer more flexibility in trading between exploration and exploitation. Nonetheless, because many of our observations could be attributed to intrinsic properties of the landscapes, we would expect many of our observations to hold in a qualitative sense for more sophisticated algorithms.

Several challenges will have to be met by developers of more sophisticated dynamic pricing techniques. It will be important to ensure that the pricing technique can track a nonstationary distribution of consumer preferences, as this is likely to drift with time. Furthermore, the dimensionality of the space being optimized may increase as attributes other than number of items are included in the pricing scheme (such as article categories, for example). We can expect this to increase the amount of time needed to learn. Perhaps the most serious challenge will be to deal with rapid shifts in the profit landscape driven by competitive interactions among two or more producers. This alone might necessitate a fundamentally new approach.

Acknowledgments

This work was supported in part by an IBM University Partnership Grant, by the IBM Institute for Advanced Commerce, and by the National Science Foundation under grant IIS-9872057.

References

- [1] P. Aghion, P. Bolton, C. Harris, and B. Jullien. Optimal learning by experimentation. *Rev. Econ. Stud.*, 58(4):621–654, 1991.
- [2] Y. Bakos and E. Brynjolfsson. Bundling information goods: Pricing, profits and efficiency. In D. Hurley, B. Kahin, and H. Varian, editors, *The Economics of Digital Information Goods*. MIT Press, Cambridge, Massachusetts, 1999.
- [3] C. H. Brooks, S. Fay, R. Das, J. K. MacKie-Mason, J. O. Kephart, and E. H. Durfee. Automated strategy searches in an electronic goods market: Learning and complex price schedules. In *Proceedings of ACM EC-99*, pages 31–40, 1999.

- [4] J. C. Chuang and M. A. Sirbu. Network delivery of information goods: Optimal pricing of articles and subscriptions. In D. Hurley, B. Kahin, and H. Varian, editors, *The Economics of Digital Information Goods*. MIT Press, Cambridge, Massachusetts, 1999.
- [5] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, 1973.
- [6] A. A. Fe'ldbaum. *Optimal Control Systems*. Academic Press, New York, 1965.
- [7] S. J. Grossman, R. E. Kihlstrom, and L. J. Mirman. A Bayesian approach to the production of information and learning by doing. *Rev. Econ. Stud.*, 44(3):533–547, 1977.
- [8] G. Keller and S. Rady. Optimal experimentation in a changing environment. *Rev. Econ. Stud.*, 66:475–507, 1999.
- [9] J. O. Kephart, R. Das, and J. K. MacKie-Mason. Two-sided learning in an agent economy for information bundles. In *Agent Mediated Electronic Commerce*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin, 2000.
- [10] J. A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.
- [11] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.
- [12] S. B. Thrun and K. Møller. Active exploration in dynamic environments. In J. Moody, S. J. Hanson, and R. P. Lippmann, editors, *Advances in Neural Information Processing Systems 4*. Morgan Kaufmann, San Mateo, CA, 1992.
- [13] R. B. Wilson. *Nonlinear Pricing*. Oxford University Press, 1993.
- [14] M. H. Wright. Direct search methods: Once scorned, now respectable. In D. F. Griths and G. A. Watson, editors, *Numerical Analysis 1995 (Proceedings of the 1995 Dundee Biennial Conference in Numerical Analysis)*, Pitma Res. Notes Math. Ser. 344, pages 191–208, Harlow, UK, 1996. Longman.

Appendix

The formula for the optimal nonlinear price schedule given by Brooks et al. [3] is incorrect. The correct form for the incremental price of the q th article is given by the following expression:

$$p_q = w \left[\frac{q(4\bar{k}N - (1 - \bar{k})q)}{4\bar{k}N} - \frac{(2\bar{k}^2N + (1 - \bar{k})^2q)\beta(q, \bar{k}, N)}{4(1 - \bar{k})^2\bar{k}N} - \frac{\bar{k}^3N\gamma(q, \bar{k}, N)}{(1 - \bar{k})^3} \right] \quad (3)$$

where

$$\beta(q, \bar{k}, N) = \sqrt{q(4\bar{k}^2N + (1 - \bar{k})^2q)} \quad (4)$$

and

$$\gamma(q, \bar{k}, N) = \log \left(1 + \frac{q(1 - \bar{k})^2 - (1 - \bar{k})\beta(q, \bar{k}, N)}{2\bar{k}^2N} \right). \quad (5)$$