Chapter 29
Pricing

Abstract The database marketing environment presents many challenges involving pricing. How should we coordinate acquisition pricing and retention pricing? How should we price when we want to re-activate customers? How should we use database marketing to price discriminate? This chapter reviews models and methods for providing insights on these questions. We point out that pricing cannot be considered in a vacuum; for example, that customer quality expectations play a key role in acquisition pricing.

Pricing is a critical area in customer management. There is a vast literature, primarily in economics, about pricing products and services. Yet, very little has been written about pricing over the lifetime of customers. This chapter will examine basic pricing theories that have applicability to pricing for individual customers over time as well as offering different customers different prices (price discrimination). A major theme is our emphasis on customer-based as opposed to product-based pricing. Customer-based pricing maximizes total customer profitability, across the several products the customer might buy, whereas product-based pricing takes the vantage point of maximizing profits for a given product. Customer-based pricing is particularly appropriate for database marketing where the focus is on managing customers.

29.1 Overview – Customer-based Pricing

Most firms price products individually. Product pricing optimizes the price of a given product/service without regard to optimizing the pricing of the total bundle of products/services the customer purchases over his or her lifetime.

Suppose a customer is more likely to buy a second product if the customer already purchased another product from the firm. Then, the firm must jointly optimize the profits based on the profitability of the two products, not each separately. An example is a financial service firm selling mortgages. Assume that the customer is more likely to purchase a mortgage if the customer has previously purchased a fee-based credit card. Hence the pricing of the fee-based credit card should reflect the purchase of future products.

Mathematically, we can set up the problem as a two-stage optimization in which the first purchase affects the probability of a second purchase. Later in the chapter we will study this structure for customer-based pricing. A related but more complex issue, not addressed currently in the marketing literature, is if the firm has two products and each could be purchased as the lead product (first product purchased), what is the optimal product-line pricing? Intuitively, both products should be priced lower.

Most of the marketing literature on acquisition pricing focuses on the pricing of durable goods with learning on the cost side and diffusion on the demand side. The issue this literature addresses is how should the firm price a new durable good overtime? Each customer is "acquired" because there is no repeat purchasing. Also, the prices for the "early adopters" affect the future price of the product through the cost side.

If the firm does not have repeat purchasing, then one commonly used solution to acquisition pricing is skim pricing – pricing high at first and then lower overtime. This is covered extensively in the literature beginning with Robinson and Lakhani (1975) and continuing to Nair (2007). Skim pricing is driven, in part, due to experience curves, if the firm has a monopoly position and customers exhibit innovator/adopter behavior. If marginal costs decline as a function of cumulative volume, then, ceteris paribus, the firm should "skim" price. If consumer tastes are different (heterogeneous) and the monopolist firm can price differentially over time, then, ceteris paribus the firm can execute a skim pricing strategy. If the discount rate for product consumption is high, then, ceteris paribus, the firm should skim price. However, if the customer is willing to wait, and the firm has a high discount rate, it may be better to price uniformly overtime or lower price. Lazear (1986) described how it is optimal to price high at the beginning of the fashion season to attract price insensitive fashion-forward consumers and then lower prices overtime to price discriminate to reach the more price sensitive customers.

Kalish (1983) examined different diffusion model assumptions. For the consumer side of the model he uses a Bass curve and for the cost side he assumes there is learning through producing. Kalish's key finding is that depending upon different assumptions, different price paths are possible. Production learning curves result in marginal costs declining over time, which always causes prices to decline. Diffusion works in the other direction causing the firm to price low to increase the number of innovators who in turn increase the number of adopters. Because diffusion and production learning curves are working in the opposite direction, the price path can decrease or increase depending upon the relative importance of the diffusion process relative to learning curves.

Numerous other articles have been written about Bass model pricing (see Bass 2004). We will not use these models in this chapter because they focus on diffusion behavior of customers. We will not assume any diffusion effects in the customer-based pricing models we will consider. Researchers may see an opportunity to incorporate diffusion in the types of customer-based pricing we will study but to do so, it would be necessary to show that there is diffusion for standard products/services that are already in the market place. We will briefly mention acquisition learning curves because many dot.com firms believed that as more customers were acquired, the cost of acquisition would decline because the firm would learn how to be more efficient in customer acquisition. However, while acquisition learning cost reductions probably exist, the dot-coms were overly optimistic. There is no documented literature showing the magnitude of these effects.

The remainder of this chapter will focus on customer-based pricing. The objective will be to maximize the long-term profitability of the customer. The reason is that customers are assets (see Blattberg et al. 2001), not simply transactions, and profit maximization must be over their long-term purchase behavior. Customer-based pricing can be broken into three components: acquisition, retention and "add-on selling" pricing. In this chapter we will focus on all three. The reason one needs to separate acquisition from retention pricing is that the price that a customer is acquired influences retention rates. Therefore, the prices are not independent. This will be discussed in detail later in the chapter. Also, because the customer's behavior is dynamic (changes over time), it is necessary to make certain assumptions about the customer's price sensitivity over time.

When one considers customer-based pricing, there are two different directions to take: (1) pricing products/services that are "one-shot" purchases but the customer purchases multiple products from the firm (e.g., electronics from Sony) and (2) pricing a repeat purchase product in which the customer purchases the same product/service multiple times. We will consider both cases in this chapter.

This chapter will be organized with Sect. 29.2 addressing customer-based pricing when customers purchase multiple one-time products from the firm; Sect. 29.3 studies pricing products/services to customers over two periods; Sect. 29.4 utilizes the customer equity model to determine acquisition and retention prices; Sect. 29.5 studies pricing to recapture customers; Sect. 29.6 addresses pricing of add-on sales; and Sect. 29.7 covers the use of price discrimination and how to value the information the firm has available to price differentially.

29.2 Customer Pricing when Customers Can Purchase Multiple One-Time Products from the Firm

Database marketers often are able to offer a "lead" product to acquire a customer. Examples include checking accounts for financial services companies, low-cost low-featured automobiles (e.g., Mercedes 1 series) for car companies, or accessories for clothing retailers. Pricing lead products is tricky because if the price is too low and repeat purchasing through other product lines is low, then the firm loses money. On the other hand, if its pricing is too high and the product quality is high, the firm may under-invest in customer acquisition.

In the model provided in this section, we will not assume the firm is price discriminating. While this is often discussed in the marketing and database marketing literature, it may be more difficult to execute in practice. Later in the chapter in the section titled Price Discrimination through Targeting Models, we cover an article by Feinberg et al. (2002) which shows some of the problems with trying to price discriminate when customers can learn about other customers' prices. The relevance of this section to database marketers is that it teaches the firm how to price "lead" products for customer acquisition. Database marketers have a major advantage in using these methods, because through their databases, they can learn which are the "lead" or first products a customer purchases.

We will begin by assuming there is a relationship between the purchase of several products because once the customer purchases one of the firm's products/services, the customer is more likely to purchase another. The products do not have to be complements. The pricing strategy and pricing levels are different than if the customer just purchased one product. For example, Dell historically sold computers but now sells printers and flat screen televisions. After purchasing one of Dell's products, its quality influences the probability of purchasing another from Dell. This should affect Dell's pricing of products that customers tend to acquire on their first purchase occasion. Dell, through its databases, can learn which are the first products customers purchase, and then consider how best to price those products recognizing their pricing affects the number of customers Dell acquires. Another example is a financial service firms that uses mortgages as its lead product and then sells traditional banking products such as checking and deposit products as second and third products.

When the products are not complements but the customer purchases multiple products from some firm, we will call these brand-linked purchases. Why is there a link? Customers learn about the brand's quality through the use of the first product. Customers have higher awareness of the brand after the initial purchase. Customers receive targeted communications from the firm. Because of the link, the firm should use customer-based pricing, not product-based pricing.

As a first example of customer-based lead product pricing, we will use a model adapted from Shapiro (1983). We will consider two products (1, 2). The firm has already set the price of product 2. The issue is: how should it price product 1 given that the probability the customer purchases product 2 depends upon the purchase and quality of product 1?

We will use an example to help make this section concrete. Suppose a firm sells two products. Product 2 is a flat screen TV, product 1 is a computer.

The firm sets the price of the flat screen television at \$2,000 and makes a margin 20% or \$400. The question is how to price the computer. The firm's cost is \$2,000. Since both products are made by the same firm, the customer estimates the quality of the flat screen television based on the quality of the computer. Both have the same brand name.

The key to this analysis is to price product 1 taking into account that after purchasing that product, the customer will revise his or her quality estimate of product 2. The assumption to be used in the model is that customers revise their quality estimate upward, although one could investigate the other case as well.

We will make the following assumptions. The firm is a monopolist who chooses product quality, q_i, for product i, which is given and not subject to change. We will use a constant cost (c_i) for the ith product. Consumers are indexed by their taste for quality for product i, θ_i, which represents the dollar value consumer of type θ_i places on product i's quality, q_i. Note that in different product categories customers may have a different preference for quality. θ_i is bounded between 0 and 1. θ_i is basically the importance of quality to the customer. A consumer of type θ_i who pays p_i for the product of quality q_i enjoys consumer surplus of $\theta_i q_i - p_i$ for $i = 1, 2$.

Consumers have an expectation about the quality of the seller's products, $R_i > 0 (i = 1, 2)$. R_i is a point expectation (not a distribution). The value of R_i is based on the firm's reputation. It does not depend upon marketing activities but could easily be related to advertising and positioning decisions by the firm.

A consumer of type θ_i will purchase initially if and only if $\theta_i R_i > p_i$ which implies $\theta_i > p_i/R_i$. Consumer diversity is captured through the distribution of θ_i across customers, denoted by $f(\theta_i)$. Then the fraction of consumers with taste parameter greater than θ_i is:

$$1 - F(\theta_i) = \int_{\theta_i}^{1} f(t)dt$$

with (29.1)

$$F(\theta_i) = \int_{0}^{\theta_i} f(t)dt$$

Demand for the initial product is given by:

$$s(p_1) = 1 - F(p_1/R_1) \tag{29.2}$$

As R_1 increases, demand increases and as p_1 increases, demand decreases. Thus, initially, if the firm could control R_1, it would try to set it as high as possible.

We will continue with our example. Suppose we assume the value of R_1 is 5,000 and θ_1 follows a uniform distribution. The firm sets a price for its computers of \$3,000. Then, the fraction of customers who will purchase the product is those for whom $\theta_1 > p_1/R_1 = 1 - (3,000/5,000) = 0.4$. Suppose the price were raised to \$4,000. Then the fraction of customers purchasing the product is $s(4,000/5,000) = (1 - F(.8)) = (1 - 0.8) = 0.2$.

We will consider two cases. Case 1: The consumer only purchases product 1. Case 2: The consumer purchases product 1 and then based on the purchase of product 1, updates the expected quality of product 2 to q_2 from R_2. The price of product 2, p_2, is set. The estimate of the quality of product 1 is R_1 and is less than the true quality q_1. However, after using product 1, the consumer updates his or her estimate of quality from R_2 to q_2.

29.2.1 Case 1: Only Product 1 Is Purchased

The customer has a purchase probability for product 1 of

$$\text{Prob(Purchase Product1)} = 1 - F\left(\frac{p_1}{R_1}\right) \tag{29.3}$$

and the profit function is:

$$\pi_1 = N\left[1 - F\left(\frac{p_1}{R_1}\right)\right](p_1 - c_1) \tag{29.4}$$

where N is the size of the market. We will assume a uniform distribution for θ_1. Then, $F(\frac{p_1}{R_1}) = \frac{p_1}{R_1}$.

Differentiating with respect to p_1 and setting the derivative equal to zero gives:

$$\begin{aligned} \frac{d\pi_1}{dp_1} &= \left(1 + \frac{c_1}{R_1}\right) - \frac{2p_1}{R_1} = 0 \ or \\ p_1 &= \frac{R_1 + c_1}{2} \end{aligned} \tag{29.5}$$

The above is based on a uniform distribution, which is tractable and shows the results clearly. Shapiro (1983) shows more general results than we show here.

Continuing with our example, let $c_1 = \$2,000$ and remembering that $R_1 = 5,000$, we have $p_1 = (5,000 + 2,000)/2 = \$3,500$. We will contrast this result to the two product purchase case.

29.2.2 Case 2: Two Product Purchase Model with Lead Product 1

In Case 2 we assume the customer purchases a second product based on his or her experience with product 1. After purchasing product 1, the customer updates his or her expectations about the quality of the firm's brand and expected quality goes from R_2 to q_2 where $q_2 > R_2$. One could also assume $R_2 > q_2$ and analyze that case. However, we will analyze the case in which consumers' expectations are below the true quality for product 2 and the consumer uses product 1 to update their expectations about the true quality of product 2, q_2.

The probability of purchasing is $1 - F(\frac{p_i}{R_i})$ for both products. Then the number of customers purchasing over both periods is:

$$S = N\left\{\left[1 - F\left(\frac{p_1}{R_1}\right)\right] + \left[1 - F\left(\frac{p_2}{q_2}\right)\right]\right\} \tag{29.6}$$

The q_2 in the second part of the equation is due to the assumption that the customer learns the true quality of product 2 after the initial purchase of product 1. Profit, π, is:

$$\pi = N\left[1 - F\left(\frac{p_1}{R_1}\right)\right] \times \left\{(p_1 - c_1) + \left[1 - F\left(\frac{p_2}{q_2}\right)\right] \times (p_2 - c_2)]\right\} \tag{29.7}$$

Let $[1 - F(\frac{p_2}{q_2})] \times (p_2 - c_2)] = k$. Again, assuming a uniform distribution for θ_1, we can optimize with respect to p_1, assuming p_2 is given. The optimal price is

$$p_1 = \frac{R_1 + (c_1 - k)}{2} \tag{29.8}$$

Because $k > 0$, the optimal price to charge for product 1 is lowered based on the additional profit to be made on product 2.

Continuing with our example, let $c_2 = \$1,000, p_2 = \$1,500$ and $q_2 = 3,000$. Then, using Equations 29.7 and 29.8 above, we have $k = (1 - F(1,500/3,000)) \times (1,500 - 1,000) = 250$. Next we substitute k into Equation 29.8 and we have $p_1 = (5,000 + (2,000 - 250))/2 = \$3,375$.

The result from the two product case shows that the profit from future purchases enters into the pricing of the first product and lowers the price when the true quality is higher than the expected quality. Thus, the firm induces more customers to sample product 1 because of its lower price which then impacts the purchase level of product 2 because the customers update their expectations about the quality of product 2. If customers do not use product 1 to update their expectations about the quality of product 2, then the firm should use the myopic price (Equation 29.5).

Obviously there are many issues associated with this model and example:

1. Why is the mean expectation for product 2 different than the true quality? When this is true, rational expectations assumptions are violated but in the real-world firms may know their product quality is greater than consumers' perceptions (through marketing research) and can influence future purchasing by using lower prices for the lead product.

2. How can a firm estimate the difference between expected quality for product 2 (R_2) and actual quality for product 2 (q_2)? This is marketing research question and can be determined by using research to ascertain the actual quality versus consumer expectations of quality. The other related issue is that once a consumer samples product 1, the firm can determine if that consumer updates his or her estimate of product 2.
3. Why is this example relevant to database marketing? Database marketers have the capability to determine which are lead products (product purchased first by customers) and can also develop targeted programs for selling the second product. Traditional marketers may also be able to use this type of pricing when products are introduced sequentially but database marketers can use their data to determine which products naturally are the lead products ever if both products are in the market.

The conclusions from analyzing these two cases are:

- Even if the lead product is a low-priced product, the firm should be very careful about its product quality because lead product quality impacts pricing and profits from future products its customers purchase firm.
- Firms need to develop mechanisms for estimating lead products and understanding how their quality levels match future products purchased by the customers.
- Firms should develop pricing policies based on the quality of its products and the purchase sequence used by customers.

29.3 Pricing the Same Products/Services to Customers over Two Periods

We now consider the case where the same product may be purchased over time, but customers do not know initial quality. The question is: should a database marketing company use a higher (lower) introductory price to acquire customers and then lower (raise) its price after acquiring the customer?

We use the same "machinery" as in Sect. 29.2. Customers are indexed by their taste for quality, θ, which represents the dollar value consumers of type θ places of the product of quality q. θ is bounded between 0 and 1. A consumer of type θ who pays p for a product of quality q enjoys consumer surplus of $\theta q - p$. Consumers have an expected quality of the seller's product, $R > 0$. R is a point expectation. The value of R is based on the firm's reputation.

A consumer of type θ will purchase initially if and only if $\theta R > p$ which implies $\theta > p/R$. Consumer diversity is captured through the distribution of θ which is denoted by $f(\theta)$. Then the number of consumers with taste parameter greater than θ is:

$$1 - F(\theta) = \int_{\theta}^{1} f(t)dt$$

or (29.9)

$$F(\theta) = 1 - \int_{\theta}^{1} f(t)dt$$

Initial demand is given by:

$$s(p) = 1 - F(p/R) \tag{29.10}$$

As R increases, demand increases. Thus, initially, if the firm could control R, it would try to set it as high as possible. However, in future periods, all consumers who learn the true quality and for whom R is greater than q, will stop buying.

Demand for the product if the quality were known (fully informed consumers) would be:

$$s(p) = 1 - F(p/q). \tag{29.11}$$

where q is substituted for R because the consumer knows the true quality. We will assume that learning occurs through personal experience and is complete and immediate.

29.3.1 Pessimistic Case: R < q – Expectations of Quality are Less than Actual Quality

The pessimistic case assumes customers have initial expectations about quality (R) below the true quality level (q). The customer, after trying the product/service learns the true quality is q. We will assume that the firm sets the introductory price (p_1) to attract customers and sets the future price for subsequent purchases (p_2).[1] We will assume a two-period model. The first-period profit function is $\pi_1 = 1 - F(\frac{p_1}{R}) \times (p_1 - c)$. Customers who try the product/service learn the true quality q. The two-period profit function is

$$\pi = \left[1 - F\left(\frac{p_1}{R}\right)\right](p_1 - c) + \left[1 - F\left(\frac{p_1}{R}\right)\right](p_2 - c) \tag{29.12}$$

The number of acquired customers is $1-F(\frac{p_1}{R})$. All of the customers who purchased in the first period will be willing to continue purchasing if $p_2/q < \theta$. The condition for these customers to purchase in the first period was $p_1/R < \theta$. Therefore all customers will be retained if $p_1/R = p_2/q$. Since $q > R$, we can make $p_2 > p_1$. Hence, because the true quality is higher than the estimated quality, the firm will raise its price in period 2.

[1] Note the subscript is now indicating period 1 versus 2, not product 1 versus product 2 as in Sect. 29.2.

Table 29.1 Comparison between two-period and one-period acquisition model

Cost (c)	Expected price (R)	Actual quality (q)	Ratio of q/R	Price in period 1 (p_1)	Price in period 2 (p_2)	Myopic price	Ratio of myopic to optimal
2	3	5	1.67	\$2.25	\$3.75	\$2.50	1.11
3	4	6	1.50	\$3.20	\$4.80	\$3.50	1.09

To determine the optimal introductory price, the firm optimizes Equation 29.12 with respect to p_1 with $p_2 = p_1(q/R)$. It can be shown that the optimal first period price is:

$$p_1 = \frac{R}{2} + \frac{c}{1 + q/R} \tag{29.13}$$

To the extent that expected quality underestimates true quality ($q > R$), we price lower in period 1 but then adjust upwards in period 2. We can contrast this with the myopic solution (one-period) which is $p_1 = (c + R)/2$ (Equation 29.5). Table 29.1 shows some results for various values of q, R and c. The two-period model chooses a lower initial than second-period price, and the myopic price is between these two.

The pricing model offered by Shapiro can be expanded to include other assumptions. For example, one can add retention to the model and discount future sales quite easily. The optimal first-period price will depend upon the level of future profits. The higher the future profits, the lower will be the price in the first period.

29.3.2 Optimistic Case: $R > q$ – Expectations of Quality are Greater than Actual Quality

When customers under estimate the true quality (pessimistic case), price should be low in the introductory period and higher in future periods. The opposite occurs when customer expectations are higher than the true quality. In this case the optimal pricing path is to charge a higher price in the introductory period and skim off those customers who incorrectly perceive quality. Then the price is reduced in the second period after the customer observes the true quality.

29.3.3 Research Issues

The work of Shapiro provides a wide variety of research questions. First, if the same price has to be offered to existing and new customers, the pricing decision becomes more complex. Database marketing potentially solves this problem because it can offer different prices to different customers. Non-database marketers have significantly greater difficulty price discriminating between new and existing purchasers.

Second, what is the impact of competition on the pricing strategy? The problem with incorporating competition is that most models assume it is a duopoly. However, in the real-world there are often many competitors and the firm is differentiated. Then, what is the impact of competition? This raises the question if it is better to assume a monopolistic model and derive the relevant pricing strategies or use a duopolist model with its attendant limitations.

A third issue for many database marketing firms is that there are many prices to be set, not just one. This is similar to a retailer's pricing problem. How does the firm develop a customer pricing strategy when it has multiple products being sold to a new customer?

A fourth issue is learning. When customers purchase a new product or service, they often are learning about the firm. Further, the firm is learning about the customer. Some papers in the economic literature on new product pricing assume the firm and/or the customer learn. Initially, the papers studied monopolistic behavior but recently have studied duopolistic behavior. See, for example, Bergemann and Valimaki (1997, 2000).

29.4 Acquisition and Retention Pricing Using the Customer Equity Model

An alternative approach to customer-based pricing is to use a customer equity model (e.g., Blattberg et al. 2001) and optimize it with respect to both acquisition and retention pricing. The Blattberg et al. (BGT) model is divided into acquisition and retention. N_t customers are available to be acquired and $N_t\alpha_t$ are actually acquired. Then, in the first period ρ_{t+1} customers are retained; in period 2 $\rho_{t+1} \times \rho_{t+2}$ customers are retained and so on. The basic model is:

$$CE(t) = \{(P_{a,t} - C_{a,t}) \times N_t\alpha_t - N_tB_{a,t}\} + \sum_{k=1}^{\infty}\left[\left\{(P_{r,t+k} - C_{r,t+k}) \times N_t\alpha_t \times r_{t+k}^k - B_{r,t+k} \times N_t\alpha_t r_{t+k-1}^k\right\}\left(\frac{1}{1+d}\right)^k\right] \quad (29.14)$$

where

$$r_{\mathrm{t+k}}^{\mathrm{k}} = \prod_{\mathrm{i=1}}^{\mathrm{k}} \rho_{\mathrm{t+i}}$$

and

$CE(t)$ = Customer equity for customers acquired at time t.
$P_{a,t}$ = Introductory price offered at time t.
$P_{r,t}$ = Retention price offered at time t.
$C_{a,t}$ = Average product cost in the acquisition period per customer at time t.
$C_{r,t}$ = Average product costs in the retention period per customer at time t.
α_t = The acquisition rate at time t.
ρ_t = The retention rate at time t.
r^k_{t+k} = The survival rate at time $t+k$.
$B_{a,t}$ = The acquisition marketing expenditures per prospect at time t.
$B_{r,t}$ = Retention marketing expenditures per existing customer at time t.
d = Discount factor
N_t = Number of prospects at time t.

The survival rate in the second term of Equation 29.14 (r^k_{t+k-1}) is lagged one period because the retention marketing expenditures are assumed to occur at the beginning of the period. We will assume that $\rho_{t+k} = \rho_t$ for all k, meaning a constant retention rate. We also will assume that the retention period price ($P_{r,t+k}$) and cost ($C_{r,t+k}$) are constant over time. These assumptions make the derivations tractable.

Pricing enters the equation through several variables in the model: acquisition and retention rates, and sales for acquisition and retention. Thus, in developing "optimal pricing," it is critical to understand the impact of price on each of these factors.

The objective function is:

$$\underset{P_a,P_r}{MAX}\ CE(t) = \left[\begin{array}{l} \{(P_a - C_a) \times (N_t \times \alpha(P_a)) - N_t \times B_a\} + \left\{ \sum\limits_{k=1}^{\infty} \{(P_r - C_r) \right. \\ \left. \times (N_t \times \alpha(P_a) \times \rho(P_a))^k \ - B_r \times N_t \times \alpha(P_a) \times \rho(P_r))^{k-1} \right\} \\ \times \left(\frac{1}{1+d}\right)^k \end{array}\right] \tag{29.15}$$

Rearranging terms in Equation 29.15 and dropping N_t yields:

$$\begin{aligned} \underset{P_aP_r}{MAX}\, CE(t) &= \left[\begin{array}{l} \{(P_a - C_a)(\alpha(P_a)) - B_a\} + \sum\limits_{k=1}^{\infty} \left[\left\{ (P_r - C_r)\alpha(P_a) \right.\right. \\ \left.\left. \times \rho(P_r)^k - B_a\alpha(P_a)\rho(P_r)^{k-1} \right\} \left(\frac{1}{1+d}\right)^k \right] \end{array}\right] \\ &= \left[\{(P_a - C_a)(\alpha(P_a)) - B_a\} + \phi\alpha(P_a) \sum_{k=1}^{\infty} \rho(P_r)^k \left(\frac{1}{1+d}\right)^k \right] \\ &= \{(P_a - C_a)(\alpha(P_a)) - B_a\} + \alpha(P_a)\phi\theta \end{aligned} \tag{29.16}$$

where

$$\phi = \frac{(\rho(P_r)(P_r - C_r) - B_r)}{1+d}, \theta = \frac{1}{1-\rho(P_r)\times r} \text{ and } r = \frac{1}{1+d}.$$

Note that $\phi\theta$ is increasing in retention rate and retention price; it therefore represents future profits after acquisition. Optimizing yields the following equations.

$$\frac{\partial \mathrm{CE(t)}}{\partial P_a} = (P_a - \mathrm{C_a}) \times \frac{\partial \alpha(P_a)}{\partial P_a} + \alpha(P_a) + \frac{\partial \alpha(P_a)}{\partial P_a} \times \phi\theta = 0 \qquad (29.17)$$

and

$$\begin{aligned} \frac{\partial CE(t)}{\partial P_r} &= \theta \frac{\partial \phi}{\partial P_r} + \phi \frac{\partial \theta}{\partial \mathrm{P_r}} = 0 \\ &=> \theta \left[(P_r - C_r) \frac{\partial \rho(P_r)}{\partial P_r} + \rho(P_r) \right] + \phi \left[\frac{r}{(1-\rho(P_r)r)^2} \frac{\partial \rho(P_r)}{\partial P_r} \right] = 0 \end{aligned} \qquad (29.18)$$

Solving Equation 29.17 for P_a reveals that the optimal acquisition price does depend upon the firm's retention tactics (i.e., retention price and marketing expenditures) and profitability. Specifically[2]:

$$\begin{aligned} P_a &= \frac{E^{\alpha}_{P_a}}{1+E^{\alpha}_{P_a}} C_a - \frac{E^{\alpha}_{P_a}}{1+E^{\alpha}_{P_a}} \phi\theta \\ &= P_m - \frac{E^{\alpha}_{P_a}}{1+E^{\alpha}_{P_a}} \phi\theta \end{aligned} \qquad (29.19)$$

where

P_m = The myopic monopolist price.
$E^{\alpha}_{P_a} = \frac{\partial \alpha(P_a)}{\partial P_a} \times \frac{P_a}{\alpha(P_a)}$ = Price elasticity of the acquisition probability.

Equation 29.19 suggests two insights concerning optimal introductory pricing:

- As future profits, represented by $\phi\theta$, become higher, the optimal acquisition price decreases.
- As future profits, represented by $\phi\theta$ approach zero, the optimal acquisition price increases towards the myopic price.

Thus customers who become more profitable over time (e.g., customers who purchase add-ons, become less costly to serve, have a strong positive influence on other customers, or have higher retention rates) should be offered a lower

[2] In a monopoly in which the firm is maximizing current period profits (i.e., behaving myopically), the optimal price for the firm to charge would be $\frac{E^{\alpha}_{P_a}}{1+E^{\alpha}_{P_a}} \times C_a$ (Pindyck and Rubinfeld 2004). Thus P_m is referred to as the myopic monopolist's price.

Table 29.2 Impact of retention price sensitivity on retention pricing

Retention spending per customer	\$0.10	–	–
Acquisition probability	10%	–	–
Discount rate	20%	–	–
Marginal cost of product	\$3.00	–	–
Model parameters	–	–	–
Exponent	1.5	–	–
Price sensitivity parameter (lambda)	0.01	0.02	0.03
Optimal retention price	\$9.53	\$7.31	\$6.39
Optimal retention rate	63.6%	54.2%	47.0%

introductory price than customers who are not as profitable in future periods. Therefore the firm should price discriminate based on the future value of a customer. This result suggests that an optimal acquisition pricing policy should be based on expected future profits because myopic pricing may not generate the optimal number of customers.[3]

Implicitly solving for the optimal retention price from the differential in Equation 29.18 is more difficult. Unlike the optimal acquisition price, the optimal retention price can not be isolated on one side of the equation and must be written as a function of the other retention parameters. The optimal retention price can be studied only through numerical examples.

We use a numerical example that depends upon the choice of the retention price model. Assume:

$$\rho(P_r) = \frac{e^{-\lambda P_r^\gamma}}{2(1+e^{-\lambda P_r^\gamma})} \tag{29.20}$$

λ was varied over three values and then an optimization routine was run to find the optimal retention price. The results are given in Table 29.2. It shows that as λ increases, making the retention rate more price sensitive, the optimal retention price decreases. The difficult practical issue is being able to determine the model's parameters. Field testing could provide data points to estimate the model.

29.5 Pricing to Recapture Customers

A related problem to acquisition and retention pricing is pricing to recapture customers. Alternative theories exist. One implies that pricing to recapture customers is based on the previous price paid, which becomes the reference price. Another states that increasing price generates a "loss" and reducing

[3] In some industries (e.g. telecommunications), this is called "sling-shot pricing." Firms price low to acquire customers, losing significant money in the first period, and then make profits through high retention rates. It is called "sling-shot pricing" because profits explode in later periods.

price is a "gain"; hence there should be asymmetry in response. The simplest hypothesis is that there is no connection between prices and reacquisition prices.

Thomas et al. (2004a) investigate various hypotheses about what affects pricing when recapturing customers. The methodology used is a split hazard model. The process begins when the customer is reacquired by the firm and ends when the customer terminates the subsequent relationship. A split hazard formulation is comprised of separate reacquisition and duration components. The reacquisition component measures the probability of recapturing a lapsed customer while the duration component predicts the length of the second tenure given the customer has been successfully recaptured.

For customer i $(i = 1 \ldots C)$ the reacquisition component is specified as a binomial probit with observation equation:

$$Z_i = \begin{cases} 1 & \text{if } \mathrm{Z_i^*} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{29.21}$$

The latent dependent variable $\mathrm{z_i}^*$ is modeled

$$\mathrm{z_i}^* = \mathrm{w_i g} + \mathrm{n_i} \tag{29.22}$$

where $w_i g$ is the deterministic component and n_i is the stochastic component; w_i is the customer's vector of predictors and g is the associated parameter vector.

Modeling of the second tenure is complicated by the firm's propensity to change the offer price during the relationship. Each consumer's second tenure consists of one or more subspells that differ only in the offer price. Assuming the termination of the relationship is independent of the current duration allows the authors to assume the duration of a subspell does not depend on the length of prior subspells. For some observations, the customer has not terminated and hence tenure is right-censored.

The model is as follows:

$$y_{is_i} = \begin{cases} y^*_{is_i} & \textit{if } \; y_{is_i} < c_{is_i} \\ c_{is_i} & \text{otherwise} \end{cases} \tag{29.23}$$

y_{is_i} is the observed duration of the relationship and c_{is_i} is the censoring value, the length that a price is offered. If the customer terminates the relationship before the price changes, then $y_{is_i} = y^*_{is_i}$. Otherwise the duration of the subspell is right censored. Latent duration $y^*_{is_i}$ is modeled as:

$$\ln(y^*_{is_i}) = X_{is_i}\beta + \varepsilon_{is_i} \tag{29.24}$$

where $X_{is_i}\beta$ is the deterministic component, ε_{is_i} is the stochastic component; X_{is_i} is the customers' vector of predictors during subspell s_i, and β is the parameter vector. Variance components are used to link acquisition and retention behavior. Customer heterogeneity is also modeled. See Thomas et al. for the specific variance structure and Chib (1993) for the estimation method.

The conclusions from Thomas et al. are:

1. Customers are more likely to be reacquired if the reacquisition price is lower.
2. The absolute effect of price is much more important than the effect of price relative to the last price paid in the prior relationship. Reacquisition strategies emphasizing decreasing price relative to the prior relationship are not likely to be effective. The most effective method for winning back lapsed customers is to offer a lower price. Also, customers who were acquired with low prices will not be enticed with significantly lower prices (in other words, relative price is not the critical factor).
3. Price increases have no effect on second tenure duration and price decreases relative to the prior price leads to a longer second tenure.
4. Higher retention prices lead to longer relationship duration. This result is different from Reinartz (2000) finding that long-life customers pay lower average prices than short-life customers. An alternative explanation is heterogeneity in reservation prices (similar to the model used by Shapiro 1983). Basically, usage allows the customer to tell true quality, which then implies a willingness to pay higher prices. Thus, the customer will pay a higher price when the customer recognizes the quality is superior through usage.

29.6 Pricing Add-on Sales

The Customer Equity model given in Equation 29.14 will be the basis for a discussion of pricing add-on sales with an additional variable added to the model.

$$\begin{aligned} CE(t) = {} & \{(P_{a,t} - C_{a,t}) * N_t\alpha_t - N_t * B_{a,t}\} + \\ & \sum_{k=1}^{\infty} \left[\left\{ \begin{array}{l} (P_{r,t+k} - C_{r,t+k}) * N_t\alpha_t{}^* \rho_{t+k}^k (\sum_{j=1}^{k-1} X_j) - B_{r,t+k} * N_t\alpha_t \\ + N_t\alpha_t{}^* \left[\rho_{t+k}^k (\sum_{j=1}^{k-1} X_j) \right] X_k(P_{ao}) m_{ao} \end{array} \right\} \left(\frac{1}{1+d}\right)^k \right] \end{aligned} \tag{29.25}$$

where

$$X_k = \begin{cases} 1 & \text{if an add-on purchase is made} \\ 0 & \text{otherwise} \end{cases}$$

m_{ao} is the margin for add-on sales.

Note that X_k depends upon the price of add-on sales but as importantly, the retention rate, ρ, depends upon the number of add-on purchases the customer makes. The model assumes that the greater the number of purchases made, the higher the retention rate. This is consistent with RFM models. In the banking industry it is believed that if the firm can sell multiple products to the customer (e.g., a mortgage, checking and savings account), the customer has more difficulty changing relationships and therefore there is greater retention. The problem is distinguishing the number of purchases (relationships) from customer heterogeneity. Many studies show that the more purchases from the firm, the greater the probability of buying. However, is this due to consumer preference for the firm's products and satisfaction with the firm or is it due to the number of purchases the customer makes?

The implications of these two possibilities are quite different. If it is the number of purchases that increases retention, then the optimal price to charge for "add-on" selling will be lower. If the cause of the higher retention rates being correlated with the number of purchases is heterogeneity in preference and satisfaction, then the firm should not lower prices when selling additional products/services.

While pricing add-on sales is a critical customer-based pricing issue, there is very little literature on the topic. One research paper that investigates a related issue is Israel (2005), who examined the relationship between recency and retention rate. Through the clever use of an automobile insurance database, Israel (2005) was able to investigate whether retention rates were affected by recency or due to unobserved heterogeneity in customer preferences. Using the auto insurance data he found evidence for both effects but concluded that the role of unobserved heterogeneity was much more important. A recent study by Reinartz et al. (2006) uses Granger causality tests to investigate whether cross-buying determines loyalty or loyalty determines cross-buying. For two firms, they conclude that loyalty causes cross-buying. If the findings from the Israel and Reinartz et al. studies were generalized, one would conclude that many of the recommendations that the number of products is a very good predictor of retention rates might in fact simply be related to consumer preference for the service. This is a crucial area for future research, and see Chapter21 for further discussion.

29.7 Price Discrimination Through Database Targeting Models

There is a vast literature in marketing and economics on price discriminating among customers. We will review several articles including one of the seminal articles on this topic by Rossi et al. (1996).

Any basic micro economics textbook discusses the conditions under which a firm can price discriminate. We will focus on how does database marketing help a firm price discriminate and can customer databases always be used to price discriminate?

We begin with an article by Rossi et al. (1996) which addresses the issue of the value of purchase history data for targeting. They compare the use of demographic data, which many marketing firms believe should serve as the basis for selecting market segments, and purchase history data.

The basic proposition of Rossi et al. is that information sets are used to draw inferences about household preferences and household-specific parameters such as price sensitivity measures. These household-specific preferences and price elasticities can be used to target promotions and prices at the individual level.

Rossi et al. compare decisions with information from the household with decisions in which there is no relevant information (causal or demographic information). The value of the information is determined by assessing how much the firm's profit increases when it uses the specific types of information versus when it does not.

The authors conclude that demographic data explains very little of the variation in price sensitivity (only 7%). Their results are consistent with other studies that find demographic information explain even less of the variability (Bucklin and Gupta (1992) and Gupta and Chintagunta (1994)). This finding is consistent with predictive models that often find demographic data have limited value in targeting. Of course, this will vary by product category and there may be instances in which demographic data could provide insights into price sensitivity (e.g., expensive durable goods).

The other conclusion Rossi et al. draw is that purchase histories have significant value even if the purchase histories are short. This is reassuring because many database marketers have limited causal data (promotional histories) captured in their databases.

The most important contribution of Rossi et al. is providing a methodology for determining the value of information to a database marketer. The methodology can be applied to situations in which the firm wants to determine the value of keeping its promotional histories for targeting. Do long purchase history data or solicitation data provide enough value to justify the cost of keeping the data?

An article by Zhang and Krishnamurthi (2004) considers the issues of when to offer promotions and how much to discount to individual customers. They use a purchase incidence, purchase timing, brand choice model which allows for inertia/variety seeking behavior and consumer heterogeneity to address these questions. Their major modeling contribution is that their model can continuously update the model's parameters by individual and then determine the optimal timing of promotions. Using their model they then evaluate different promotional strategies. Their finding is that by varying the timing of promotions based on continuously updating the model, profits can increase.

One issue Zhang and Krishnamurthi raise is consumers anticipating promotions (consumer expectations). Others such as Erdem et al. (2003) and Hendel and Nevo (2002) have developed models that include consumers' expectations about prices. This is a critical issue and was discussed earlier in this chapter. If consumers anticipate price changes, then the models need to be altered to take into account how changes in promotional or price strategies lead to changes in consumers' expectations which then change consumers' purchase patterns. This phenomenon is well understood for durable goods (particularly electronics and computers) where consumers know that prices will drop in the future. For database marketers, as pricing strategies become dynamic, consumer expectations may also be dynamic and understanding the implications of having both changing at once becomes an important modeling challenge. Research has just scratched the surface of this issue.

In another paper, Zhang and Wedel (2007) ask the question about what level of personalized promotions work in different environments. They consider two types of retailers – bricks and mortar (offline) and online – and three levels of promotions – mass market, segment level and individual-level. They build a similar model to that of Zhang and Krishnamurthi which includes a purchase incidence, choice and quantity model and allows for heterogeneity.

One interesting finding by Zhang and Wedel (2007) is that the gains in profits from mass market to segment to individual-level optimization are not that large, especially in off-line environments, while there is a substantial gain in moving from current managerial practice to mass market optimization. This is a provocative result. One factor driving their results is the retailer's cost structure. For offline retailers the ability to reach its customers through mass-communications may be more cost-effective than it is for online retailers because they have the store environment to communicate promotions, which is very low cost. Also, the cost of offering promotions through tools such as Catalina Marketing Systems is low for offline retailers. Hence the gains from personalization are offset by the low delivery costs. Another factor is that both offline and online retailers have data on their customer's behavior with *their* business, but not with competitors. This might add significant measurement error to individual-level and even segment-level price elasticity estimates (see Park and Fader 2004).

Zhang and Wedel do indicate that individual level databases do have the potential to provide significant economic returns in the correct environment. However, for a database marketer when the costs of mass promotions are very low, the payouts from having an extensive customer database may also be very low. This is exemplified by the struggles offline grocery retailers worldwide are having with the use of their customer databases. With the exception of a few retailers such as Tesco (UK) or CVS (USA), very few have been able to generate a significant payout from frequent shopper programs. This may be caused by the fact that mass promotions are more cost effective.

One final issue associated with price discrimination is do consumers/ customers care if others are receiving a lower price than they do? Feinberg et al. (2002) address this issue using laboratory experiments. Their basic premise is that consumer preference for a firm's products is affected not just by the price the consumer pays but the prices that other consumers are paying.

To test this premise, they design a series of laboratory experiments in combination with various models to answer the question above. They find two effects beyond traditional economic rationality: (1) consumers feel *betrayed* if their favored firm offers promotions to switchers; and (2) consumers are *jealous* if another firm offers a price decrease to its own loyal customers and the favored firm does not.

These findings are in a laboratory setting in which the authors can manipulate the information each consumer receives. In the real-world, it is not clear how much of this information is available to customers. However, regarding finding one above, consumers can often tell if a firm is offering lower prices to new customers. Mobile phone firms offer lower prices to new customers; credit card companies offer low teaser rates to prospective new customers. This type of offer may lead to a sense of betrayal and therefore affect the ability of the firm to offer low prices to new customers and then raise their prices over time. The findings of Feinberg et al. are provocative and should lead to additional research, particularly understanding in the real-world the prevalence of consumers' knowledge about what others pay and what their reaction is to that information (see Krishna et al. (2004) for an interesting follow-up paper on whether betrayal and jealousy effects exist when firms target price *increases* selectively).