

# UC Berkeley

## CUDARE Working Papers

### **Title**

Learning, Forgetting, and Sales

### **Permalink**

<https://escholarship.org/uc/item/7td275hv>

### **Authors**

Villas-Boas, Sofia B

Villas-Boas, Miguel

### **Publication Date**

2006-06-01

# LEARNING, FORGETTING, AND SALES\*

SOFIA BERTO VILLAS-BOAS  
*(University of California, Berkeley)*

J. MIGUEL VILLAS-BOAS  
*(University of California, Berkeley)*

June, 2006

---

\* Address for correspondence: Haas School of Business, University of California at Berkeley, Berkeley, CA 94720-1900. E-mail: sberto@are.berkeley.edu and villas@haas.berkeley.edu.

# LEARNING, FORGETTING, AND SALES

## ABSTRACT

Sellers of almost any product or service rarely keep their prices constant through time, and frequently offer price discounts or sales. This paper investigates an explanation of sales as a way for uninformed consumers to be willing to experience the product, and learn about its fit, and where informed consumers may forget about (or change) their preferences. We investigate the role of the rate of forgetting on the timing between sales, and of the rate of learning and menu costs on the length of a sale. We also investigate the effect of a seller carrying multiple products on the pattern of sales. Using price series from supermarket categories, and given the assumed simplified preference structure, we obtain empirical estimates of the rates of learning and forgetting, and of the other model parameters.

## 1. INTRODUCTION

Sellers of almost any product or service rarely keep their prices constant through time, and frequently offer price discounts or sales. For example, Zhao (2006) documents empirically the existence of significant temporal price variation in several product categories in the grocery channel. Practitioners often justify the offering of sales as a way to “stimulate early purchase” (e.g., Kotler and Keller, 2006, p. 452), re-draw attention to a product, or allow the consumer to learn “about the performance of the brand” (Blattberg and Neslin, 1990, p. 118). In fact, some authors go on to claim that “most promotional pricing ... is intended to induce buyers to try a product” (Nagle, 1987, p. 196), and past research has often distinguished between trial and post-trial purchases (e.g., Silk and Urban, 1978).

This paper examines this explanation by considering a model where consumers can be uncertain about their valuation for a product, and are only willing to try it, and find out about how much the product fits their preferences, if the product is priced at a sufficiently low level - the sale. The consumers that learn that their product fit is high are then willing to pay more for the product, and the firm may then be able to charge a high (regular) price and get these informed consumers to buy the product. The possibility of consumers learning about the product gives then incentives for a firm to temporarily cut its price in order to induce the consumers to try and learn about the product.

However, through time consumers may forget about how much they value the product. With the passing of time the number of consumers that have forgotten about how much they value the product may become so high that again it pays off for the firm, to cut its price to induce these consumers to try the product again, and be reminded of their valuation for the product. The existence of consumer forgetting (in addition to learning) can then lead to temporary sales / price cuts that are repeated through time after periods where the firm charges high (regular) prices.<sup>1</sup>

The rate at which consumers forget about their product valuation determines then the rate at which the number of uninformed consumers increases, and therefore when the firm is again tempted to lower the price to induce the uninformed consumers to re-try the product. Given the rate of

---

<sup>1</sup>Equivalently to forgetting, some consumers may leave the market and other consumers may enter the market who are uninformed about their product valuations. Because in many markets where sales are observed we may not see a substantial number of consumers leaving or coming into the market, we keep the “forgetting” interpretation throughout the paper. Another equivalent interpretation is some consumers changing preferences, with the change of preferences not allowing them to figure out if the product is a good fit to the new preferences. These alternative interpretations are further discussed below in greater detail. For further discussion on consumer forgetting see, for example, Keon (1980), and for some empirical evidence see Mehta et al. (2004), and the papers discussed there.

forgetting one can then obtain what is the optimal time interval between successive sales, with higher rates of forgetting leading to more frequent sales.

Another important dimension to consider is that in most markets sales last for longer than the minimum amount of time necessary for a firm to cut its price and then raise it back again. For example, in the grocery channel, a sale / price cut does not last for only one day or one week, but typically lasts for two or four weeks. This means that if this learning and forgetting features are a reasonable explanation of sales then they must also provide an explanation for the length of a sale. In fact, if both consumer instantaneous learning about their valuation for a product is not possible and there are menu costs of changing prices, it turns out that a sale needs to last longer than the minimum amount of time necessary for a firm to cut its price, for the firm to build sufficiently up the stock of consumers that are informed about their product valuation. Then, the rate at which consumers learn about the product valuation can determine the time interval during which the product is on sale, with faster learning leading to shorter time intervals when a sale is in place.

Another interesting aspect to investigate is what happens when a seller carries more than one product, and whether it should stagger the sales of the products sold, or offer all sales simultaneously. We find that in this case, when there are some consumers that have high preference for some brands and some consumers that are indifferent between the brands, there is a force towards offering sales simultaneously. This allows the seller not to offer too many sales for the consumers that are relatively indifferent across brands.

Note also that given price series data one can infer the market characteristics discussed above. In fact, given the length of the sales periods and the interval between sales, one can make inferences, given the preference model, about the rates of learning and forgetting in the market. Using price series from supermarket categories, we obtain some estimates for these parameters.

Several theories have been presented to explain sales. Given demand uncertainty, if firms make orders before demand is realized, then they may have to cut prices to move inventory if realized demand falls short of expectations (e.g., Lazear 1986, Pashigian 1988), or when learning about demand (e.g., Aghion et al., 1991). Other authors have argued that firms may use sales to price discriminate between high valuation-high inventory costs and low valuation-low inventory costs consumers (e.g., Blattberg et al. 1981, Jeuland and Narasimhan 1985).<sup>2</sup> Another explanation that has been presented relies on different price information by consumers, or a discrete number of segments with different preferences across products, to generate mixed price strategy equilibria in

---

<sup>2</sup>See, for example, Macé and Neslin (2004) for some empirical evidence of the effects of such consumer behavior at the time of promotions.

competition, which can be interpreted as sales (e.g., Shilony 1977, Varian 1980, Rosenthal 1980).<sup>3</sup> Another possibility is that if competing firms do not change prices simultaneously, one may be able to obtain Edgeworth price cycles in equilibrium (e.g., Maskin and Tirole, 1988). Another explanation relying on competition, is when the repeated interaction among firms allows for collusive outcomes to emerge, and demand changes through time affecting what is possible to collude on (e.g., Green and Porter 1984, Rotemberg and Saloner 1986). There is also a literature explaining sales in durable goods through the existence of generations of consumers coming into the market every period, and where sales are offered to sell to the accumulated stock of low valuation consumers who wait for a sale (e.g., Conlisk, Gerstner, and Sobel 1984, Sobel 1991), and a literature on competition with homogeneous switching costs, where equilibrium stochastically time dependent prices can also be interpreted as sales (e.g., Padilla 1995, Anderson et al. 2004).

The most related paper to the analysis presented here is Bergemann and Valimaki (BV, 2006). BV considers monopoly pricing with experience goods. BV distinguishes between two situations: (1) One is when the uninformed consumers expected valuation is low compared to the optimal monopoly price to the informed consumers, such that prices start out low and rise through time. (2) The other is when the uninformed consumers expected valuation is high compared to the optimal monopoly price, such that prices start out high and decline through time. More related to this paper, Section 6 of BV, considers the case where at each moment in time some consumers leave the market and are substituted by new consumers (which is an equivalent interpretation to learning and forgetting, as discussed above). BV shows that in case (1) above, the steady-state equilibrium involves the firm charging a high price with some probability (the price only for the informed consumers), and a low price with the complementary probability (the price that also attracts the uninformed consumers). As noted in BV, the fact that the low price is charged with some probability can be interpreted as a sale to attract the uninformed consumers. As also noted there, the equilibrium with a probability of charging a low price results from the continuous time formulation used in that paper, and, in fact, the equilibria become in pure strategies in discrete time, with a low price charged if there are not “enough” informed consumers, and a high price charged otherwise. This paper concentrates on case (1) above, the uninformed consumers expected valuation is low compared to the optimal monopoly price to the informed consumers. First, the paper considers a simple two-period overlapping generations model to re-derive the result that temporary sales are offered (price cycles). Second, the paper considers a longer time horizon in discrete time with immediate learning

---

<sup>3</sup>See also Narasimhan (1988a), Raju et al. (1990), and Lal and Villas-Boas (1998). An empirical test can be found in Villas-Boas (1995). For an empirical analysis of the independence through time of price promotions see Rao et al. (1995). See also Rao (1991) for a model distinguishing regular prices and price promotions, and Anderson and Simester (1998) for a model presenting a role for sale signs.

after the first experience, to focus on the properties of the interval of time between successive sales (related to the probability of sale above). Third, the paper considers a continuous time version with gradual learning (as considered in BV) and adds menu costs of changing prices, to generate the result that the period of time when the seller offers sales (or charges a high price) lasts for some time. As discussed above, this fits reality better, as sales are typically offered for some length of time, and not only for the minimum time period needed to cut and raise prices. We then confront the model with price series data from the grocery channel, and obtain parameter estimates, in particular, estimates on the rates of learning and forgetting. Finally, the paper also presents results for the case in which the seller carries more than one product.

Three other papers are closely related to the results presented here: Doyle (1986), Narasimhan (1988b), and Freimer and Horsky (2003). Doyle considers a free-entry competition model with some consumers leaving, and some consumers entering the market in every period, where new (firms) entrants offer and advertise low prices, and then keep the consumers that found a good fit in future periods. In contrast, this paper can be seen as arguing that competition or advertising are not necessary for the existence of sales when consumers learn about their preferences for the products through experience. Narsimhan presents a model where consumers only try the product if the price is low enough, but once having tried the product, consumers are willing to purchase at a higher price. “New triers” decline due to attrition, and the model results in temporary sales and some of the price dynamics shown here. Freimer and Horsky argue that consumer learning may lead to the existence of sales with a model of lagged purchase effects, and show properties of the demand function such that temporary sales are optimal. These properties of the demand function are not directly linked in that paper to what consumers learn through experience.

The remaining of the paper is organized as follows. The next section presents a simple model illustrating that sales can exist in equilibrium with consumer learning and forgetting. Section 3 looks at the optimal interval between sales, Section 4 considers the issue of the optimal length of a sale, and Section 5 presents an empirical application of the model to supermarket data, obtaining some estimates of the model parameters. Section 6 presents an extension to the case when firms sell multiple products. Finally, Section 7 concludes.

## 2. A SIMPLE MODEL OF SALES WITH LEARNING AND FORGETTING

Consider a monopolist selling a non-durable non-storable product in every period of an infinite number of periods to a market of mass one. In each period the monopolist sets a price and consumers decide whether to buy or not, given the information that they have. The marginal cost of production

is zero. A fraction  $\alpha$  of consumers derives utility  $u$  of consuming the product, while a fraction  $1 - \alpha$  derives utility zero. Consumers can either be informed or not informed about their utility for the product. They are informed if they consumed the product in the previous period and did not forget about that information since then. They are uninformed if they either did not consume the product in last period, or consumed it and forgot about the information since then. Consumers forget their information every two periods, one half in the odd periods, one half in the even periods.<sup>4</sup> Consumers are assumed to be risk neutral. We are looking for the Markov Perfect equilibria, that is, the subgame perfect equilibria where the players' actions only depend on the payoff-relevant state variables.

With this demand structure the optimal price for the informed consumers is  $u$ , while the price to attract the uninformed consumers is strictly less than  $u$ . This is the case in BV where the uninformed consumers expected valuation,  $\alpha u$ , is low compared to the optimal monopoly price to the informed consumers,  $u$ .<sup>5</sup> In this section we are looking for an equilibrium, where the seller alternates between selling to the uninformed consumers, and selling only to the informed consumers. Because the optimal price when selling only to the informed consumers is  $u$ , the uninformed consumers when deciding whether to buy the product know that, even if they have a good experience, in the next period they will have zero surplus. That means that the uninformed consumers only buy if the price is below or equal to  $\alpha u$ , and, therefore, the optimal price when attracting the uninformed consumers is  $\alpha u$ . The equilibrium would then be with alternating prices between the regular price  $u$  and the sale price  $\alpha u$ .

Let us now investigate whether there are no profitable deviations from this possible equilibrium. Denote by  $V(x)$  the net present value of profits for the monopolist if there is a fraction  $x$  of consumers informed about the product among the consumers that did not forget any information that they might have from the last period. Consider a period where in the beginning of the period no consumer is informed about the product. Then, if the firm charges a price equal to  $\alpha u$ , it gets a net present value of profits equal to  $\alpha u + \delta V(1)$ , while if it charges any price strictly above  $\alpha u$ , no consumer buys this period and the firm gets a net present value of profits equal to  $0 + \delta V(0)$ , where  $\delta$  is the discount factor. Charging a price of  $\alpha u$  is better if  $\alpha u + \delta[V(1) - V(0)] \geq 0$ .

---

<sup>4</sup>As noted above, this could be equally interpreted as a model of overlapping generations of consumers, each living for two periods. In the next sections, we consider the case where the consumers forgetting their information are just a fraction of the informed consumers.

<sup>5</sup>Note that in a two-point distribution of the consumer experiences we could still get the other case in BV, the uninformed consumers expected valuation being high compared to the optimal monopoly price for the informed consumers. To see this, just make the poor experience different from zero, say  $\underline{u}$ . Then, if  $\underline{u}$  is high enough (but below  $u$ ) the optimal monopoly price for the informed consumers is  $\underline{u}$ , which is lower than the uninformed consumers expected valuation,  $\alpha u + (1 - \alpha)\underline{u}$ .



Consider now a period where in the beginning of the period all the consumers who did not forget any information that they might have (one half of the market) are informed about the product. Then, if the firm charges a price of  $u$ , it gets a demand in that period of the consumers that did not forget and had a good experience with the product,  $\alpha/2$ . The net present value of profits would then be  $\alpha u/2 + \delta V(0)$ . If the firm deviated and charged a price of  $\alpha u$ , it would get a demand in that period of the consumers that did not forget and had a good experience with the product,  $\alpha/2$ , plus the consumers that forgot any prior information that they had,  $1/2$ , for a total demand of  $(1 + \alpha)/2$ . The net present value of profits would then be  $\alpha(1 + \alpha)u/2 + \delta V(1)$ . Charging a price of  $u$  is better if  $\alpha^2 u/2 + \delta[V(1) - V(0)] < 0$ .

Given the candidate equilibrium strategies, price of  $u$  if  $x = 1$ , and price of  $\alpha u$  if  $x = 0$ , we have  $V(0) = \alpha u \frac{1+\delta/2}{1-\delta^2}$  and  $V(1) = \alpha u \frac{1/2+\delta}{1-\delta^2}$ . Checking the no deviation conditions above, one can then obtain that no deviations are profitable from the candidate equilibrium if  $\alpha \leq \frac{\delta}{1+\delta}$ . That is, under this condition we get that the monopolist optimal behavior involves sales every two periods at the low price,  $\alpha u$ . In the next section we show that this equilibrium with successive temporary sales is the natural equilibrium to consider when consumers forget at random time periods (and possibly less frequently on average).

### 3. THE TIME INTERVAL BETWEEN SALES

In this section we discuss the question of the optimal time interval between sales. In order to study this question we now consider a variation of the model above where in each period a fraction  $1 - \beta$  of the consumers forget about any information that they might have. The idea is that if the rate of forgetting,  $1 - \beta$ , is small enough, the seller might prefer to keep the price high longer (as compared to just one period in the section above), as there are few consumers to inform about the product.

Consider first the price to be charged such that only the informed consumers buy the product. In that case the maximum that can be charged is  $u$ , and any informed consumer buys as long as the price is below or equal to  $u$ . Then, the optimal price to be charged when selling only to the informed consumers is  $u$ .

Consider now the price to be charged such that the uninformed consumers are also willing to buy the product. Denote it by  $p$ . Note first that after  $p$  is charged, all consumers become informed about the product. Therefore, because we are restricting attention to Markov perfect equilibria, the sequence of prices that follow the price  $p$ , is always the same, and independent of the history

up to the price  $p$  being charged. This also implies that the highest price  $p$  at which the uninformed consumers are willing to buy is independent of the history of the market.

The equilibrium will then involve the firm charging the price  $p$  every  $T$  periods, with the price  $u$  charged in between attracting only the informed consumers that value the product. That is, after all consumers are informed, the seller charges a price of  $u$  during  $T - 1$  periods, after which it charges the price  $p$ , all consumers become informed again, and the cycle re-starts. The number of periods  $T$  between sales is the object of this section.

Before analyzing  $T$ , note first that the highest price  $p$  at which the uninformed consumers are willing to buy is such that the expected net present value of utilities of an uninformed consumer buying the product is zero. After paying the price  $p$ , the consumer can expect to obtain with probability  $\alpha$  a utility of  $u$  today, plus a sequence every  $T$  periods of  $u - p$  for as long as the consumer does not forget his information. Note that when the seller charges a price of  $u$  the consumer has zero surplus. Formally, the condition for  $p$  is then  $\alpha u - p + \alpha(u - p)(\delta\beta)^T \frac{1}{1 - (\delta\beta)^T} \geq 0$ , which results in

$$p = \frac{\alpha u}{1 - (1 - \alpha)\delta^T \beta^T}. \quad (1)$$

The sale price  $p$  which attracts the uninformed consumers is greater than the expected utility of consumption in a period,  $\alpha u$ , as, once informed, the consumer has the potential of getting a positive surplus in future periods when the seller offer again a sale. This gain of being informed is greater the shorter the time interval to the next sale, the lower the rate of forgetting, and the greater the discount factor  $\delta$ .

Consider now the equilibrium interval between sales,  $T$ . Let  $x_t$  be the fraction of consumers that is informed at time  $t$ , and  $V(x_t)$  the net present value of profits at time  $t$ , if the fraction of consumers informed at time  $t$  is  $x_t$ .

Consider a time period  $t$  where the firm is considering between offering the sale in that period, or waiting one period, and offering the sale in the next period. By offering the sale in period  $t$ , the net present value of profits is  $p(\alpha x_t + 1 - x_t) + \delta V(\beta)$ , as the fraction of informed consumers in period  $t + 1$  will be  $\beta$ , after all consumers become informed in period  $t$ . By delaying the sale for one period, the net present value of profits in period  $t$  is  $\alpha x_t u + \delta p(\alpha \beta x_t + 1 - \beta x_t) + \delta^2 V(\beta)$ . By equalizing these two terms one can obtain the level of the state variable  $x_t$  such that the firm chooses to offer the sale. Except for integer issues related to  $T$ , one can obtain the equilibrium interval between sales by making  $x_t = \beta^T$  in that equality, i.e.,

$$\alpha \beta^T u + \delta p(\alpha \beta^{T+1} + 1 - \beta^{T+1}) = p(\alpha \beta^T + 1 - \beta^T) + \delta(1 - \delta)V(\beta). \quad (2)$$

Finally, to complete the derivation of the time interval between sales  $T$  we need to obtain  $V(\beta)$ , the net present value of profits after a period in which a sale was offered. Because, the time interval between sales is  $T$  we can have  $V(\beta) = \alpha\beta u + \delta\alpha\beta^2 u + \delta^2\alpha\beta^3 u + \dots + \delta^{T-2}\alpha\beta^{T-1}u + \delta^{T-1}p(\alpha\beta^T + 1 - \beta^T) + \delta^T V(\beta)$  from which one can obtain

$$V(\beta) = \frac{1}{1 - \delta^T} \left[ \alpha\beta u \frac{1 - (\delta\beta)^{T-1}}{1 - \delta\beta} + \delta^{T-1} p (\alpha\beta^T + 1 - \beta^T) \right]. \quad (3)$$

Putting together (1), (2), and (3) one can obtain the equilibrium time interval between sales  $T$  and do comparative statics of the model parameters. This statement of the market equilibrium and the comparative statics results are presented in the following proposition (proofs in the Appendix).

**PROPOSITION 1:** *The market equilibrium involves sales every  $T$  periods, where  $T$  is obtained from (1), (2), and (3). For the discount factor  $\delta$  close to one and  $\beta$  large, the time interval between sales  $T$  is greater, the lower the rate of forgetting  $1 - \beta$ , the lower the probability of product fit  $\alpha$ , and the lower the discount factor  $\delta$ .*

The lower the rate of forgetting, the more consumers remain informed through time, and, therefore, the less need there is to offer a sale for the uninformed consumers to be willing to try the product. The greater the probability of product fit, the greater the share of the informed consumers who can buy the product, leading to a greater incentive to offer a sale. A greater discount factor yields less relative benefits of offering a sale as the present value of the future profits are not too hurt by delaying the sale.

In order to have a sense of the equilibrium number of periods between sales, consider pricing decisions per week, a yearly interest rate of 4%, leading to  $\delta = .999$  (per week), a rate of forgetting of 1% per week (a yearly forgetting rate of 40%), leading to  $\beta = .99$ , and a probability of product fit  $\alpha = 10\%$ . This would then lead to an equilibrium number of weeks between sales of nine weeks. Figure 1 presents the evolution through time of prices, unit sales, and the fraction of informed consumers for this set of parameter values. Table 1 presents the equilibrium number of weeks between sales for different values of  $\beta$ .

Note that the results above present the equilibrium time interval between sales when the seller cannot commit to a sequence of prices. The seller, when making a decision of whether or not to offer the sale, has to take as given its, and the consumers', best-response behavior in the future. That is, the seller has to take as given the highest price  $p$  needed to attract the informed consumers to try the product. This is represented by condition (2), which can be seen as representing the maximum of  $V(\beta)$  in (3) over  $T$ , while taking  $p$  as fixed. The optimum commitment interval between sales

Table 1  
Time Interval Between Sales as a  
Function of the Rate of Forgetting

$\beta$	yearly rate of forgetting	$p/u$	$T$ (weeks)
.97	80%	.39	6
.975	73%	.43	6
.98	64%	.45	7
.985	55%	.48	8
.99	41%	.55	9
.995	23%	.62	12
.999	5%	.75	22

The Table presents the value  $T$  for different values of  $\beta$ , for  $\delta = .999$  and  $\alpha = 10\%$ .

would result from the maximization of (3) while taking into account that  $p$  is a function of  $T$ . Since  $p$  is decreasing in  $T$ , that is, the price to attract the uninformed consumers has to be lower if the possibilities of getting a lower price in the future are less frequent, the optimal commitment interval between sales is lower than the non-commitment interval between sales. This result is intuitive: in the non-commitment case consumers are concerned about trying the product because they will be charged a high price later on, and, ex-post, it is in the best interest of the seller to charge such high price. This result of the commitment time interval between sales being smaller than the one in equilibrium can be seen as similar to the result in BV that the probability of offering a sale is greater under commitment than under non-commitment. The comparative statics results above of the equilibrium time interval between sales with respect to the rate of forgetting and the discount factor can also be seen as similar to the same comparative statics of the probability of offering a sale in BV, with the probability of offering a sale in that paper increasing in the discount factor and in the rate of forgetting.

#### 4. THE DURATION OF A SALE

In the previous section sales lasted only for one period, or equivalently, for the shortest amount of time that the seller can have between lowering and raising back again the price. However, in the real world sales last longer than the minimum amount of time needed for two successive price changes. In order to study this question of the duration of a sale, we restrict attention to continuous time (in order to get sharper results), and introduce gradual learning, and menu costs of changing prices. For time periods when price is high, the equation describing the evolution of the set of

informed consumers in continuous time is determined, as above, by a constant rate of forgetting among informed consumers, to obtain  $\frac{dx}{dt} = \dot{x} = -(1 - \beta)x$ , where  $\beta < 1$  (and  $\beta$  can be negative). For time periods when the price is low, attracting uninformed consumers, the equation describing the evolution of the set of informed consumers under gradual learning (as a constant fraction of the set of uninformed consumers) is  $\dot{x} = \lambda(1 - x) - (1 - \beta)x$ , where  $\lambda$  represents the learning parameter. When  $\lambda$  approaches infinity, learning is infinitely fast, and we are back in the model of the previous section where the main issue was the time interval between sales. This model of gradual learning is the same as the one considered in BV. Gradual learning could be justified either because consumers only feel the need for the product at some moments in time (and they only learn when they feel the need, and experiment the product), or because consumers do not always learn about their fit when they experiment the product.

In this model, in order to get the result of a positive duration of a sale one needs both gradual learning and menu costs of changing prices. With positive menu costs but without gradual learning, the optimal strategy for the seller would be to simply cut and raise prices infinitely fast, when offering a sale. With gradual learning but without menu costs, the optimal strategy for the seller in steady state is to offer a sale at each moment in time with some probability (“chattering,” the term used in the context of advertising pulsing, e.g., Mahajan and Muller 1986, Villas-Boas 1993), as the decrease in the measure of informed consumers is greater just after the sale (proportional to the measure of informed consumers), and the increase in the measure of informed consumers is greater when the sale starts (proportional to the measure of uninformed consumers). This is the case considered in BV. Below we briefly discuss other possible model formulations without menu costs that would yield a positive duration of sales.

Note first that the uninformed consumers at the end of the sale period are only willing to purchase the product if the sale price is  $\alpha u$ , as the probability of them learning their valuation approaches zero. This means that the sale price has to be  $\alpha u$ . Let  $k/2$  be the costs incurred each time the seller changes its price. Let  $r$  be the continuous time interest rate. In an optimal strategy after charging the low price for some time, the threshold fraction of informed consumer when the seller switches to the high price is always the same. Let it be  $x_h$ . From this point on the market is in steady-state with cycles of fixed length, let it be  $T$ . In the first part of the cycle after  $x_h$  the seller charges the high price  $u$ . Let us denote the fraction of the cycle when the high price is charged as  $z$ , such that the first part of the cycle has length  $zT$ . In the second part of the cycle the seller charges the low price  $\alpha u$ . This second part of the cycle lasts for  $(1 - z)T$  time periods.

The net present value of profits after  $x_h$  is reached and the price has changed to the high price

is

$$V(x_h) = \frac{1}{1 - e^{-rT}} \left\{ \int_0^{zT} \alpha u x e^{-rt} dt - \frac{k}{2} e^{-rzT} + \int_{zT}^{(1-z)T} \alpha u (\alpha x + 1 - x) dt - \frac{k}{2} e^{-rT} \right\} \quad (4)$$

where  $\dot{x} = -(1 - \beta)x$  for  $0 \leq t < zT$ , and  $\dot{x} = \lambda(1 - x) - (1 - \beta)x$  for  $zT \leq t < T$ , where the time  $t$  is zero at the beginning of the cycle considered. From this we obtain  $x = x_h e^{-(1-\beta)t}$  for  $0 \leq t < zT$ , and  $x = (x_h e^{-(1-\beta)zT} - \frac{\lambda}{1+\lambda-\beta}) e^{-(1+\lambda-\beta)(t-zT)} + \frac{\lambda}{1+\lambda-\beta}$  for  $zT \leq t < T$ .

Because we know that at the end of the cycle the fraction of the informed consumers is the same as at the beginning of the cycle, we can obtain

$$x_h = \frac{\lambda}{1 + \lambda - \beta} \frac{1 - e^{-(1+\lambda-\beta)(1-z)T}}{1 - e^{-(1+\lambda-\beta)T}}, \quad (5)$$

from which one can obtain for  $z$  and  $\beta$  close to one that  $x_h$  is decreasing in  $T$ ,  $z$ , and increasing in  $\lambda$  and  $\beta$ . The result on  $T$  comes from, as discussed above, considering learning proportional to the measure of uninformed consumers, and forgetting proportional to the measure of informed consumers, yielding shorter cycles to be better. The results on  $z$ ,  $\beta$ , and  $\lambda$  are straightforward: longer periods with a high price, more forgetting, and slower learning all lead to less informed consumers.

Maximizing  $V(x_h)$  with respect to  $z$  and  $T$  one obtains the optimal cycle length, and the optimal duration of the sale period,  $(1 - z)T$ . The following result presents some of the comparative statics with respect to the optimal  $z$  and  $T$ .

**PROPOSITION 2:** *The optimal duration of the cycle  $T$  is increasing in the menu costs  $k$ , and when  $k \rightarrow 0$ , the optimal duration of the cycle  $T$  converges to zero, but more slowly than  $k$ . When  $k \rightarrow 0$ , the fraction  $z$  of the duration of the cycle with the high price converges to  $\frac{1}{\lambda}(2 - \beta - \sqrt{(1 - \beta)^2 + \frac{(1-\beta)(1+\lambda-\beta)}{1-\alpha}})$ . For  $k$  small we have that the optimal fraction  $z$  of the duration of the cycle with the high price is decreasing in the forgetting rate  $(1 - \beta)$ , the learning rate  $\lambda$ , and the probability of good fit  $\alpha$ . For  $k$  small and the optimal  $z$  close to one, the optimal duration of the cycle  $T$  is decreasing in the forgetting rate  $(1 - \beta)$ , the learning rate  $\lambda$ , the probability of good fit  $\alpha$ , the utility of good fit  $u$ , and the interest rate  $r$ .*

The effect of menu costs  $k$  on the optimal duration of the cycle  $T$  results from the savings of menu costs of changing prices less often. When the menu costs get closer to zero, the optimal duration of the cycle approaches zero, approaches “chattering” as argued above. Consider now the comparative statics on the fraction  $z$  of time of the cycle when the high price is charged. Intuitively, the more consumers forget the less appealing it is for the firms to keep the high prices, as the number of informed consumers decreases faster. Interestingly, the faster consumers learn, the more it pays

for the firm to keep the price low in order for more consumers to have an opportunity to learn about their fit. Finally, the greater the probability of fit, the less the price needs to be cut to attract the uninformed consumers, and, therefore, it is more appealing for the seller to keep the lower price for a longer period. These comparative statics on  $z$  are similar to the ones in BV for the probability of offering a sale.

Consider now the comparative statics on the optimal duration of the cycle. The more consumers forget the more important it is to cut the prices to attract the uninformed consumers, leading to a shorter duration of the cycle. Similarly, the faster consumers learn, the faster the number of informed consumers increases when the price is low, leading the firm to raise the price sooner, a shorter duration of the cycle. The greater the probability of good fit and the greater the utility, the bigger the size of the market, and, therefore, in relative terms it is like the menu costs are lower, leading to a shorter duration of the cycle. Finally, the more patient (lower  $r$ ) the seller is, the more it is willing to have a longer cycle.

Combining the effects on  $z$  and  $T$  we can see the effects of the different variables on the duration of the sale  $(1 - z)T$ .

**PROPOSITION 3:** *Consider the menu cost  $k$  parameter small. Then, for the optimal  $z$  close to one, the optimal duration of a sale  $(1 - z)T$  is increasing in the forgetting rate  $(1 - \beta)$ , the learning rate  $\lambda$ , and the probability of good fit  $\alpha$ , and decreasing in the utility of good fit  $u$ , and the interest rate  $r$ .*

When the forgetting rate increases, the number of uninformed consumers decreases faster, and therefore, the seller has to offer a longer sale period in order to replenish the stock of informed consumers. At the same times, when either the learning rate or the probability of good fit increases, the seller finds it more appealing to extend the length of the sale period, as there is a greater payoff in terms of an increased number of informed consumers, or it is less costly in terms of loss in revenue of offering a price that attracts the uninformed consumers. When the size of the market increases (represented by the utility of good fit  $u$ ), it becomes more important for the firm to extract the surplus of the informed consumers, and the duration of the sale period is shortened. Finally, when the seller is more patient, it values more the future gains from the informed consumers, and chooses to offer a longer sale period.

In order to have a sense of the implications of these results in a market setting, consider the case where the yearly interest rate is 4% (resulting in a yearly continuous interest rate of  $r = 3.9\%$ ), a  $\beta = .9$  (resulting in about 1.7% rate of forgetting per week, a maximum percentage of informed

Table 2  
Cycle and Sale Duration in Continuous Time Model

$1 - \beta$	$\lambda = 2$		$\lambda = 3$	
	cycle days	sale days	cycle days	sale days
.04	193	25	533	59
.06	139	22	186	24
.08	117	21	139	21
.10	106	21	117	20

The Table presents the cycle and sale durations ( $T$  and  $(1 - z)T$ , respectively) in days, for different values of  $\beta$  and  $\lambda$ , given  $k = 1.04$ ,  $u = 4600$ ,  $r = .039$ , and  $\alpha = 10\%$ .

consumers if the low price is charged forever of  $\frac{\lambda}{1+\lambda-\beta} = 96\%$ , a probability of good fit of  $\alpha = 10\%$ , menu costs of changing prices of \$.52 per price change (resulting in  $k = \$1.04$ ), and .7% of the maximum potential revenues under three sales per year (resulting in  $3k = .007\alpha u \frac{\lambda}{1+\lambda-\beta}$ ).<sup>6</sup> This then results in optimal cycles of about 117 days, and optimal sales durations of about 20 days. Figure 2 presents the evolution through time of price, of unit sales, and of the fraction of informed consumers for this set of parameter values (up to scale of the price and unit sales series). Table 2 presents the optimal cycle and sale durations for different values of the forgetting and learning rates.

Let us now briefly discuss other possible model formulations that would yield a positive duration of sales. For this positive duration to obtain without menu costs one needs, (1) for the high price, for the measure of informed consumers to decrease less just after the sale price is offered than after some time after the sale price is offered, and/or, (2) for the sale price, for the measure of informed consumers to increase less just when the sale price starts to be offered than when the sale price is in place after some time. Point (1) could happen, for example, if consumers are less likely to forget just after they learn about the product fit, or if consumer can learn not only through experience, but also through word-of-mouth of the informed consumers (in a non-linear way). The former would require considering more state variables in the model, rather than just the measure of informed consumers. The latter brings in a non-linearity that may be seen as relatively orthogonal to the object of study. Although both effects may be present in real markets, and worth exploring in future research, modelling them can add substantial complications to the analysis.<sup>7</sup>

---

<sup>6</sup>These values of menu costs of \$.52 per price change, and .7% of revenues are the estimates presented in Levy et al. (1997) for the supermarket industry.

<sup>7</sup>See, for example, Feinberg (2001), for a discussion of related issues in the context of dynamic advertising policies.



## 5. EVIDENCE AND PARAMETER CALIBRATION IN THE SUPERMARKET INDUSTRY

The model above is relatively parsimonious, while seeming to describe the successive sales behavior in the supermarket industry. In order to have some sense of the parameter values in the real world, we performed the parameter estimation for several price series of the leading brands of several product categories in the supermarket industry.

We considered the pricing behavior during two years of the upc item with the largest market share in each of several product categories in a store (belonging to a large supermarket chain) in a U.S. metropolitan area in the early 90's. The product categories considered are Bacon, Tissue, Butter, Margarine, and Coffee. Given the weekly price series for each of the upc's, we constructed for each upc a series of successive interval time lengths alternating between regular and sale price. We then considered an observation (a cycle) as an interval time length where the regular price was charged followed by the interval time length where the sale price was charged. For each upc each observation would then have a length of the cycle,  $T_i$ , and the fraction of the cycle where the regular price was charged,  $z_i$ , where  $i$  represents an observation.

From the maximization of  $V(x_h)$  above with respect to  $T$  and  $z$  we get the optimal cycle length and fraction of the cycle with the high price as a function of the parameters,  $T^*(\beta, \lambda, \alpha, k, u, r)$  and  $z^*(\beta, \lambda, \alpha, k, u, r)$ . The observations are interpreted as being equal to these optimal values plus an error term,  $T_i = T^*(\beta, \lambda, \alpha, k, u, r) + \varepsilon_{T_i}$  and  $z_i = z^*(\beta, \lambda, \alpha, k, u, r) + \varepsilon_{z_i}$ . This is a simple econometric model that allows us to illustrate the parameters of the model above in real markets, for the purposes of this paper. We discuss below possible richer versions of the econometric model that could make it a better description of reality, and their complications. The error terms in the model considered could potentially be interpreted as decision errors.

For each product we estimate two parameters through the method of moments with two moment conditions, with the other parameters calibrated using other information. The moment conditions used are one for the cycle length,  $E[T_i - T^*(\beta, \lambda, \alpha, k, u, r)] = 0$ , and one for the fraction of the cycle when the regular price is charged,  $E[z_i - z^*(\beta, \lambda, \alpha, k, u, r)] = 0$ . As  $T^*(\cdot)$  and  $z^*(\cdot)$  do not change across observations in the assumed model, the two moment conditions allow us to only identify two parameters.

Assuming a yearly interest rate of 4%, we get a continuous interest rate  $r = 3.9\%$ . From Levy et al. (1997), we have an estimate of menu costs of \$.52 per price change for the supermarket industry, yielding  $k = 1.04$ . In order to get an estimate of  $\alpha$ , suppose that in terms of the model above consumers with probability  $\alpha$  value the product at some high valuation, which will be the

Table 3  
Some Parameter Estimates in Supermarket Industry

Parameters	Bacon	Tissue	Butter	Margarine	Coffee
$1 - \beta$	.53 (.18)	.15 (.10)	1.01 (.00)	.73 (.11)	1.01 (.00)
$\lambda$	25.9 (10.3)	26.8 (9.7)	12.8 (2.0)	15.3 (7.1)	19.5 (3.9)
$\alpha$	.50	.59	.47	.28	.33
$u$	20245.0	1699.1	2297.2	3692.6	3158.3
Cycle days	55 (7.3)	99 (18.3)	37 (1.8)	37 (3.7)	37 (3.6)
$z$	.81 (.04)	.89 (.02)	.66 (.02)	.78 (.03)	.76 (.02)

The Table presents parameter estimates for data for leading products in several categories in a supermarket in a U.S. major metropolitan area. Standard errors are in parentheses.

regular price, and with probability  $(1 - \alpha)$  value it at marginal cost. Then  $\alpha$  represents both the fraction of the informed consumers that buy the product at the regular price and the fraction of the margin of the regular price that corresponds to the margin at the sale price. Denoting the regular price as  $p_R$ , the sale price as  $p_S$ , and the marginal cost as  $c$ , we would then have  $\alpha(p_R - c) = p_S - c$  from which we can get  $\alpha = (\frac{p_S}{p_R} + m_R - 1)/m_R$  where  $m_R$  is the percentage margin,  $\frac{p_R - c}{p_R}$ . Using the percentage margin  $m_R$  as 50% (this includes both retailer and manufacturer margin),<sup>8</sup> and the average ratio of sale price to regular price for a product, we get an estimate of  $\alpha$  for that product. Finally, to get an estimate of  $u$  given the other parameters, we considered the estimate of profits in a week after a sale concluded as  $m_R$  times revenues to be equal to the model profits as  $\alpha x^h u \frac{7}{365}$ , where  $m_R$  was set at 50% as above,  $x^h$  is the fraction of consumers informed about the product after a sale is concluded, obtained from (5), and  $\frac{7}{365}$  is just the fraction of a year corresponding to a week.

With these constraints on the parameters, we then estimated the following remaining parameters: the parameter associated with forgetting,  $\beta$ , and the parameter associated with learning,  $\lambda$ . The estimates for these parameters for each of the five products, as well, as the corresponding estimates for the other products are presented in Table 3 (standard errors are computed through bootstrap).

According to the model, from the estimates consumers seem to be more forgetful in Butter and

---

<sup>8</sup>This 50% margin estimate is in the region of estimates from, for example, Villas-Boas (2005) and Villas-Boas and Zhao (2005), and roughly consistent with the information obtained from interviews with managers at manufacturers and retailers.

Coffee, and less so in Bacon and Tissue. Consumers seem also to learn faster in Bacon and Tissue, and to learn more slowly in Butter.

Obviously these estimates have to be interpreted with care. The model is rather parsimonious, able to recover the patterns in the data, and relatively easy to work with to get parameter estimates. However, the econometric model includes decision errors that were not considered in the theoretical model, and some of the model assumptions may be seen as not too realistic. We now discuss some of these issues.

In the interest of parsimony, the model above does not include any dimension of uncertainty that could potentially lead to variable patterns of cycles, sale periods, prices, and quantities demanded, which we observe in the data. For empirical purposes it would be interesting to extend the model to allow for dimensions of uncertainty that would allow for variable patterns in the observed variables, rather than rely on “decision errors”. This exercise would be interesting, but at the same time would require a numerical analysis of the firm’s stochastic dynamic optimization, beyond the purpose of this paper. It would also be interesting to make further use of the price and quantity information in the data, in addition to the limited use presented above.

The consumer preference heterogeneity and learning in the model above is also relatively stylized in order to illustrate the market effects in the simplest possible way. However, in an empirical application it could also make sense to have a richer consumer preference heterogeneity, and allowing for a more flexible consumer learning structure. In the supermarket industry firms carry several substitute products, and it could also be interesting in a detailed empirical application to consider this possibility, while the model above only considered the firm selling one product. In the next section we briefly consider the potential effects of the firm carrying more than one product. Other potential issues to consider involve competition and the channel structure in the industry.

## 6. THE MULTIPRODUCT SELLER

In most cases sellers carry more than one product, and, therefore, it is important to investigate what happens in such a situation. This section briefly considers such a case, when the memory of consumers lasts only for two periods, as in Section 2, and the seller carries two products. The main question that this section tries to address is whether the seller prefers to offer sales in both products simultaneously, or offer sales in a staggered way. In order to address this question consider the model of Section 2, with a market of mass two (mass one “forgetting in every period”), where the probability of fit for each product is  $\alpha$  and the probability of fit is independent across products.<sup>9</sup>

---

<sup>9</sup>The case of the correlation of fit across products is discussed below.

Because of its importance in the optimal sale policy we introduce further a segment of consumers of mass  $\gamma$  who know that they have a good fit with both products, and that just buy at the lowest price. The case of  $\gamma = 0$  is just a particular case.

If consumers receive the same surplus from buying either product, they buy the product for which they know that they have a good fit, if they are informed about that product, and if uninformed about both products, they buy either product with equal probability. Let the two products be products 1 and 2, and  $x_i$  be the measure of consumers that is informed about product  $i$  at the beginning of the period. The state variable of the problem can then be seen as  $(x_1, x_2)$ , and let  $V(x_1, x_2)$  be the net present value of the profits of the seller if starting the period in state  $(x_1, x_2)$ .

Consider the state when all consumers that did not forget are informed about product 1,  $x_1 = 1, x_2 = 0$ . Then if the firm charges a price of  $u$  for product 1 and  $\alpha u$  for product 2, then it gets a demand of  $\alpha$  for product 1, and a demand of  $2 - \alpha + \gamma$  for product 2, for a payoff of  $\alpha u + \alpha u(2 - \alpha + \gamma) + \delta V(0, 1)$  as, next period, all consumers that did not forget would be informed about product 2. Similarly, if the seller charged a high price  $u$  for both products, the payoff would be  $(\alpha + \gamma)u + \delta V(0, 0)$ , while if the seller charged a low price  $\alpha u$  for both products, the payoff would be  $\alpha u(\alpha + \frac{1}{2} + \frac{\gamma}{2}) + \alpha u(\frac{3}{2} - \alpha + \frac{\gamma}{2}) + \delta V(\frac{1}{2}, \frac{1}{2})$ .

Consider now the state where one half the consumers that did not forget are informed about one product, and the other half are informed about the other product,  $x_1 = x_2 = \frac{1}{2}$ . Then, if the seller charges a high price  $u$  on both products it gets a payoff of  $(\alpha + \gamma)u + \delta V(0, 0)$ . On the other hand, if the seller only charges a high price on one of the products, it gets a payoff of  $\frac{\alpha}{2}u + \alpha u(1 + \frac{1-\alpha}{2} + \gamma) + \delta V(1, 0)$  (assuming symmetry,  $V(1, 0) = V(0, 1)$ ). Finally, consider the state where no consumer is informed,  $x_1 = x_2 = 0$ . Then, if the seller offers the sale price  $\alpha u$  on both products it gets a payoff of  $\alpha u(2 + \gamma) + \delta V(\frac{1}{2}, \frac{1}{2})$ . Offering the sale price on only one product yields a payoff of  $\alpha u(2 + \gamma) + \delta V(1, 0)$ , and charging a high price in both products yields  $\gamma u + \delta V(0, 0)$ .

Consider the two alternative steady states: the one where the product offered on sale varies from period to period, and the one where both products are offered simultaneously on sale every two periods. In the steady-state where the product offered on sale varies from period to period, the profit per period is  $\alpha u(3 - \alpha + \gamma)$ . Alternatively, in the steady-state where both products are offered simultaneously on sale every two periods, the average profit across periods is  $\frac{u}{2}[3\alpha + \gamma(1 + \alpha)]$ . Note that the former is greater than the latter if and only if  $\alpha(3 - 2\alpha) - \gamma(1 - \alpha) > 0$ . That is, if the probability of fit  $\alpha$  is sufficiently small, or if the measure of consumers that always like both products  $\gamma$  is sufficiently large, then the simultaneous offer of sales leads to higher average profits per period. Alternatively, if the measure of consumers that always like both products is small (in

the extreme case,  $\gamma = 0$ ), then the staggered offer of sales leads to higher average profits per period. For a sufficiently patient seller ( $\delta$  high enough) this comparison is the relevant one to figure out the optimal sales policy, and is presented in the following proposition.

**PROPOSITION 4:** *Suppose that the seller is sufficiently patient, and that  $\gamma(1 - \alpha) < 3\alpha$ . Then the steady-state has sales in each product offered every two periods. The steady state has staggered (simultaneous) sales on both products if  $\alpha(3 - 2\alpha) - \gamma(1 - \alpha) > (<)0$ .*

In order to gain intuition on this result, consider first the case where  $\gamma$  is small. In that case the seller prefers to offer staggered sales as this allows the consumers who had a poor experience with one product in one period to still be offered a sufficiently low price on the other product in the following period, so that they are willing to buy the product. However, when  $\gamma$  is large, by offering staggered sales the seller allows for a substantial number of consumers willing to buy at the high price, to buy at the low price in every period. In this case, by offering simultaneous sales, the seller is able to charge this set of consumers a high price at least every two periods.<sup>10</sup>

It is also interesting to consider what happens if there is some positive correlation across products of the product-fit. In that case, a staggered sales offering would allow consumers that have a good fit on one product to buy the other product on sale, as they are aware of the positive correlation of fit across products. In order to stop this behavior the firm would then be better off in offering simultaneous sales on both products. That is, the positive correlation across products of the product-fit is a force towards offering simultaneous sales.

## 7. CONCLUDING REMARKS

This paper argues that consumer learning and forgetting can generate successive sales in equilibrium. The rate of consumer forgetting affects the time interval between sales, and the duration of the sale can be affected by both the rate of learning and the menu costs of changing prices. We calibrated the model parameters for data from the supermarket industry (given the restrictive assumptions on consumer preferences), and presented results for the case in which the seller carries more than one product.

There are several aspects of the general problem that deserve attention in future research. First, it would be interesting to have a more general model of consumer preferences, and multiproduct

---

<sup>10</sup>The condition  $\gamma(1 - \alpha) < 3\alpha$  is just to guarantee that selling only to the consumers who always like both products (with measure  $\gamma$ ) is never the optimal policy.

pricing to look more carefully at the empirical calibration of the model. Second, it would be interesting to investigate what happens in competition. Some of the insights in that case may be similar to the case of competition with switching costs (as in Padilla 1995), but there may be other insights, given the consumer learning and willingness to try different products only if the price is low enough. Finally, it would be interesting to research the implications on distribution channel contracting of this consumer learning behavior leading to sales.

## APPENDIX

PROOF OF PROPOSITION 1: Using (1) and (3) in (2) and making  $\delta \rightarrow 1$  one obtains

$$T(1 - \beta)\beta^T[1 + (1 - \alpha)(1 - \beta - \beta^T)] - (1 - \beta^T)(1 - (1 - \alpha)\beta^T) = 0. \quad (\text{i})$$

Noting that the left hand side is decreasing in  $T$  and  $\alpha$ , and increasing in  $\beta$  for  $\beta$  large, one obtains that  $T$  is increasing in  $\beta$  and decreasing in  $\alpha$ . Finally, the equivalent to (i) with  $\delta \neq 1$ , is increasing in  $\delta$  for  $\delta$  close to one, obtaining that  $T$  is increasing in  $\delta$ .

PROOF OF PROPOSITION 2: Because the derivative of  $V(x_h)$  with respect to  $k$  is increasing in  $T$  and independent in  $z$  we know that that  $T$  is increasing in  $k$ . Differentiating  $V(x_h)$  with respect to  $T$ , multiplying by  $(1 - e^{-rT})$ , and equalizing to zero, one obtains:

$$\begin{aligned} & -re^{-rT}V(x_h) + \alpha ux_h ze^{-(1+r-\beta)zT} + \alpha u \left[1 - \frac{\lambda(1-\alpha)}{1+\lambda-\beta}\right] (e^{-rT} - ze^{-rzT}) \\ & - \alpha(1-\alpha)u(x_h e^{-(1-\beta)zT} - \frac{\lambda}{1+\lambda-\beta}) (e^{-(1+r+\lambda-\beta)T} - ze^{-(1+r+\lambda-\beta)zT} + \frac{rk}{2}(ze^{-rzT} - e^{-rT})) = 0. \quad (\text{ii}) \end{aligned}$$

From this one can obtain that when  $k \rightarrow 0$  we have  $T \rightarrow 0$ . One can then obtain,

$$\lim_{k \rightarrow 0} V(x_h) = \alpha ux_h z + \alpha u(1-z) \left[1 - \frac{\lambda(1-\alpha)}{1+\lambda-\beta}\right] - (1-\alpha)\alpha u \left(x_h - \frac{\lambda}{1+\lambda-\beta}\right) (1-z) \quad (\text{iii})$$

and  $\lim_{k \rightarrow 0} x_h = \frac{\lambda(1-z)}{1+\lambda(1-z)-\beta}$ . In order to find the optimal  $z$  for  $k$  close to zero, one can maximize (iii) to obtain

$$z = \frac{1}{\lambda} \left(2 - \beta - \sqrt{(1-\beta)^2 + \frac{(1-\beta)(1+\lambda-\beta)}{1-\alpha}}\right). \quad (\text{iv})$$

Differentiating (iv) with respect to  $\beta$ ,  $\alpha$ , and  $\lambda$  one obtains that  $\frac{dz}{d\beta} > 0$ ,  $\frac{dz}{d\alpha} < 0$ , and  $\frac{dz}{d\lambda} < 0$ .

In order to consider the comparative statics with respect to  $T$  for  $k$  small, multiply (ii) by  $(1 - e^{-rT})$  and divide by  $T^2$  to obtain

$$\begin{aligned} \lim_{k \rightarrow 0} \frac{2k}{z(1-z)\alpha u T^2} &= \lambda \frac{z(1+r-\beta) - r}{1+\lambda(1-z)-\beta} + r \left[1 - \frac{\lambda(1-\alpha)}{1+\lambda-\beta}\right] \\ &+ [(1+r+\lambda-\beta)(1+z) - r] \frac{\lambda(1-\alpha)(1-\beta)}{(1+\lambda-\beta)[1+\lambda(1-z)-\beta]}. \quad (\text{v}) \end{aligned}$$

Differentiating (v) for  $z$  in (iv) close to one we can obtain  $\frac{dT}{d\beta} > 0$ ,  $\frac{dT}{d\alpha} < 0$ ,  $\frac{dT}{d\lambda} < 0$ ,  $\frac{dT}{du} < 0$ , and  $\frac{dT}{dr} < 0$ .

PROOF OF PROPOSITION 3: Differentiating the duration of the sale  $(1 - z)T$  with respect to a variable  $y$  one obtains  $\frac{d(1-z)T}{dy} = -T\frac{dz}{dy} + (1 - z)\frac{dT}{dy}$ . For  $k$  close to zero, we can obtain from (v)  $\frac{dT}{dy} = \frac{T}{2(1-z)}\frac{dz}{dy} - \frac{T^3(1-z)}{4k}\frac{d}{dy}\left(\frac{2k}{T^2(1-z)}\right)$ . This then yields that for  $k$  close to zero, and the optimal  $z$  close to one, the sign of  $\frac{d(1-z)T}{dy}$  is the same as the sign of  $-\frac{dz}{dy}$  for  $y \in \{\beta, \lambda, \alpha\}$ , and the same as the sign of  $\frac{dT}{dy}$  for  $y \in \{u, r\}$  (given the assumption of  $k$  converging to zero faster than  $z$  converging to one).



## REFERENCES

- AGHION, P., P. BOLTON, C. HARRIS, AND B. JULLIEN (1991), "Optimal Leaning by Experimentation," *Review of Economic Studies*, **58**, 621-654.
- ANDERSON, E.T., AND D.I. SIMESTER (1998), "The Role of Sale Signs," *Marketing Science*, **17**, 139-155.
- ANDERSON, E.T., N. KUMAR, AND S. RAJIV (2004), "A Comment on : 'Revisiting Dynamic Duopoly with Consumer Switching Costs' ," *Journal of Economic Theory*, **116**, 177-186.
- BERGEMANN, D., AND J. VALIMAKI (2006), "Monopoly Pricing of Experience Goods," *Journal of Political Economy*, forthcoming.
- BLATTBERG, R.C., G.D. EPPEN, AND J. LIEBERMAN (1981), "A Theoretical and Empirical Investigation of Price Deals for Consumer Nondurables," *Journal of Marketing*, **45**, 116-129.
- BLATTBERG, R.C., AND S.A. NESLIN (1990), *Sales Promotion – Concepts, Methods, and Strategies*, Prentice Hall: Englewood Cliffs, New Jersey.
- CONLISK, J., E. GERSTNER, AND J. SOBEL (1984), "Cyclic Pricing by a Durable Goods Monopolist," *Quarterly Journal of Economics*, **93**, 489-505.
- DOYLE, C. (1986), "Intertemporal Price Discrimination, Uncertainty and Introductory Offers," *The Economic Journal*, **96**, 71-82.
- FEINBERG, F. (2001), "On Continuous-Time Optimal Advertising under S-Shaped Response," *Management Science*, **47**, 1476-1487.
- FREIMER, M., AND D. HORSKY (2003), "Try It You Will Like It – Does Consumer Learning Lead to Competitive Price Promotions?," *working paper*, University of Rochester.
- GREEN, E., AND R. PORTER (1984), "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, **52**, 87-100.
- JEULAND, A.P., AND C. NARASIMHAN (1985), "Dealing–Temporal Price Cuts – by Seller as a Buyer Discrimination Mechanism," *Journal of Business*, **58**, 295-308.
- KEON, J. (1980), "The Bargain Value Model and a Comparison of Managerial Implications with the Linear Learning Model," *Management Science*, **26**, 1117-1130.
- KOTLER, P., AND K.L. KELLER (2006), *Marketing Management*, Pearson Prentice Hall: Upper Saddle River, New Jersey.
- LAL, R., AND J.M. VILLAS-BOAS (1998), "Price Promotions and Trade Deals with Multi–Product Retailers," *Management Science*, **44**, 935-949.
- LAZEAR, E. (1986), "Retail Pricing and Clearance Sales," *American Economic Review*, **76**, 14-32.

- LEVY, D., M. BERGEN, S. DUTTA, AND R. VENABLE (1997), "The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains," *Quarterly Journal of Economics*, **112**, 793-825.
- MACÉ, S., AND S.A. NESLIN (2004), "The Determinants of Pre- and Postpromotion Dips in Sales of Frequently Purchased Goods," *Journal of Marketing Research*, **41**, 339-350.
- MAHANJAN, V., AND E. MULLER (1986), "Advertising Pulsing Policies for Generating Awareness for New Products," *Marketing Science*, **5**, 89-106.
- MASKIN, E., AND J. TIROLE (1988), "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles," *Econometrica*, **56**, 571-599.
- MEHTA, N., S. RAJIV, AND K. SRINIVASAN (2004), "Role of Forgetting in Memory-based Choice Decisions," *Quantitative Marketing and Economics*, **2**, 107-140.
- NAGLE, T. (1987), *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, Prentice Hall: Englewood Cliffs, New Jersey.
- NARASIMHAN, C. (1988a), "Competitive Promotional Strategies," *Journal of Business*, **61**, 427-449.
- NARSIMHAN, C. (1988b), *A Model of Discounting for Repeat Sales*, in T.M. Devinney, ed., "Issues in Pricing – Theory and Research," Lexington Books.
- PADILLA, A.J. (1995), "Revisiting Dynamic Duopoly with Consumer Switching Costs," *Journal of Economic Theory*, **67**, 520-530.
- PASHIGIAN, B.P. (1988), "Demand Uncertainty and Sales: A Study of Fashion and Markdown Pricing," *American Economic Review*, **78**, 936-953.
- RAJU, J.S., V. SRINIVASAN, AND R. LAL (1990), "The Effects of Brand Loyalty of Competitive Price Promotional Strategies," *Management Science*, **36**, 276-304.
- RAO, R.C. (1991), "Pricing and Promotions in Asymmetric Duopolies," *Marketing Science*, **10**, 131-144.
- RAO, R.C., R. ARJUNJI, AND B.P.S. MURTHI (1995), "Game Theory and Empirical Generalizations Concerning Competitive Promotions," *Marketing Science*, **14**, G89-100.
- ROSENTHAL, R. (1980), "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price," *Econometrica*, **48**, 1575-79.
- ROTEMBERG, J., AND G. SALONER (1986), "A Supergame-Theoretical Model of Business Cycles and Price Wars During Booms," *American Economic Review*, **76**, 390-407.
- SHILONY, Y. (1977), "Mixed Pricing in Oligopoly," *Journal of Economic Theory*, **14**, 373-388.
- SILK, A.J., AND G.L. URBAN (1978), "Pre-Test Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," *Journal of Marketing Research*, **15**, 171-191.

- SOBEL, J. (1991), "Durable Goods Monopoly with Entry of New Consumers," *Econometrica*, **59**, 1455-1485.
- VARIAN, H. (1980), "A Model of Sales," *American Economic Review*, **70**, 651-659.
- VILLAS-BOAS, J.M. (1993), "Predicting Advertising Pulsing Policies in an Oligopoly: A Model and Empirical Test," *Marketing Science*, **12**, 88-102.
- VILLAS-BOAS, J.M. (1995), "Models of Competitive Price Promotions: Some Empirical Evidence from the Coffee and Saltine Crackers Markets," *Journal of Economics and Management Strategy*, **4**, 85-107.
- VILLAS-BOAS, J.M., AND Y. ZHAO (2005), "Retailer, Manufacturers, and Individual Consumers: Modeling the Supply Side in the Ketchup Marketplace," *Journal of Marketing Research*, **42**, 83-95.
- VILLAS-BOAS, S.B. (2005), "Vertical Relationships Between Manufacturers and Retailers: Inference With Limited Data," *working paper*, University of California, Berkeley.
- ZHAO, Y. (2006), "Price Dispersion in the Grocery Market," *Journal of Business*, forthcoming.

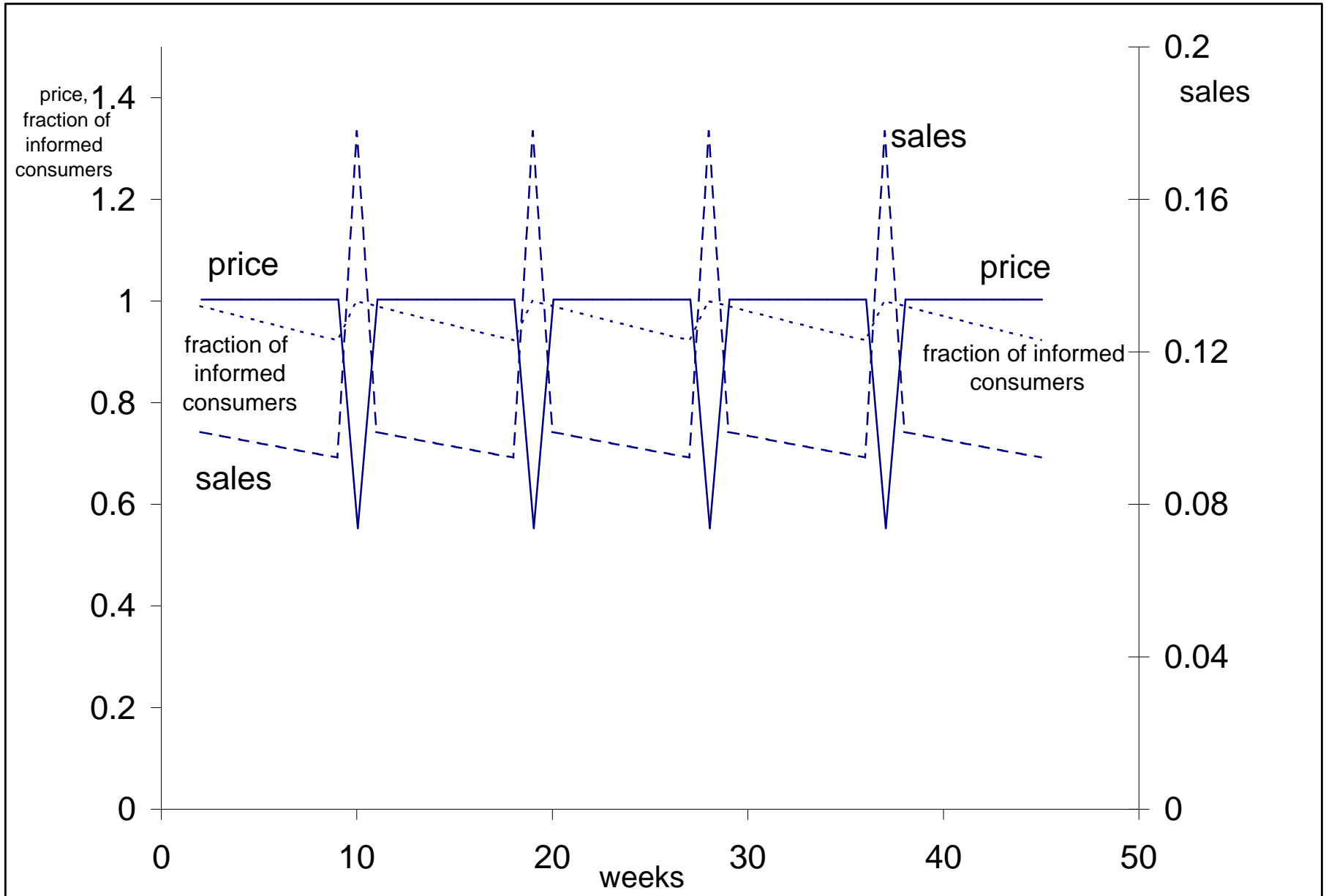


Figure 1: Evolution through time of prices, unit sales, and fraction of informed consumers for discrete model.

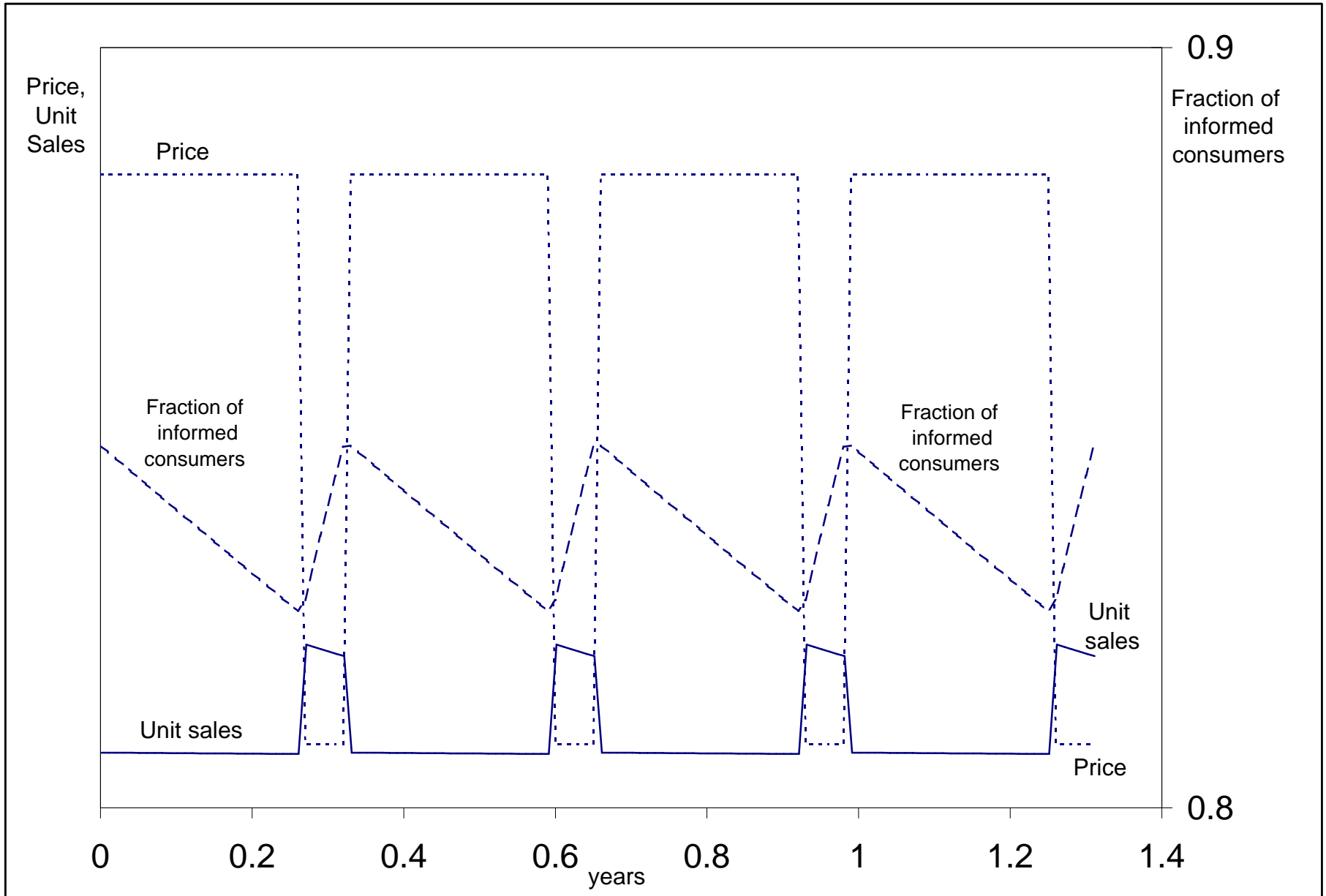


Figure 2: Evolution through time of price, unit sales, and fraction of informed consumers for continuous time model with gradual learning.