Title
Learning, Forgetting, and Sales

Permalink

Journal

Authors
Villas-Boas, Sofia B
Villas-Boas, J Miguel

Publication Date
2008

Peer reviewed

Learning, Forgetting, and Sales

Sofia Berto Villas-Boas
Department of Agricultural and Resource Economics, University of California at Berkeley,
Berkeley, California 94720, sberto@are.berkeley.edu

J. Miguel Villas-Boas
Haas School of Business, University of California at Berkeley, Berkeley, California 94720,
villas@haas.berkeley.edu

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 consumer forgetting on the timing between sales, and of the rate of consumer learning and menu costs on the length of a sale. The rate of consumer forgetting can be linked to the length of purchase cycle and the level of consumer involvement. We show that the discount frequency and the discount depth are increasing in the rate of consumer forgetting, and that the discount frequency is increasing in the learning rate. The duration of a sale is increasing in the rate of consumer forgetting and the rate of consumer learning.

Key words: marketing; promotion; sales; learning and forgetting
History: Accepted by Jagmohan S. Raju, marketing; received May 21, 2007. This paper was with the authors
    3 months for 3 revisions. Published online in *Articles in Advance* September 5, 2008.

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" (Kotler and Keller 2006, p. 452), redraw 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 then

gives incentives for a firm to temporarily cut its price to induce the consumers to try and learn about the product. However, over time consumers may forget about how much they value the product. Consumers can be seen as more likely to forget about the value of the product for categories with longer purchase cycles, for products with which consumers have less involvement, or for products with greater complexity. With the passing of time the number of consumers that have forgotten about how much they value the product may become so high that it again 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.[1] The rate at which consumers forget about their product valuation then determines

[1] 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

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 retry the product. Given the rate of 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 consumer 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 and raise it back again, for the firm to build up sufficiently the stock of consumers who 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. In Villas-Boas and Villas-Boas (2006)—an earlier version of this paper—we argue 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 toward 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.[2]

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).[3] 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 competition, which can be interpreted as sales (e.g., Shilony 1977, Varian 1980, Rosenthal 1980).[4] 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 et al. 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). Neither of these explanations, although capturing important market situations, corresponds to the anecdotal explanation provided by practitioners that is described above, of offering price promotions to induce trial. Furthermore, there may be market situations in grocery markets where none of these explanations may apply, or may be empirically important.[5]

The paper most related to the analysis presented here is Bergemann and Välimäki (2006). Bergemann and Välimäki (2006) consider monopoly pricing with experience goods in continuous time and distinguish between two situations: (1) 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) 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, §VI of Bergemann and Välimäki (2006) considers the case where at each moment in time some consumers leave the market and are replaced by new consumers (which is an

---

[2] In Villas-Boas and Villas-Boas (2006) we argue that, 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. There, we also present some estimates for supermarket categories of the rates of learning and forgetting using such inferences.

[3] See, for example, Macé and Neslin (2004) for some empirical evidence of the effects of such consumer behavior at the time of promotions.

[4] 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, Anderson and Simester (1998) for a model presenting a role for sale signs, and Kuksov (2004) for the potential effects of search costs on the consumers' knowledge of prices charged.

[5] It would be interesting to empirically investigate the relative relevance of the different explanations for sales.

consumer forgetting, see, for example, Keon (1980), and for some empirical evidence see Mehta et al. (2004), and the papers discussed there.

Table 1   Summary of Relation of the Paper with Existing Literature

	Price promotions interpretation	Strategic consumers	Time interval between sales	Duration of a sale	Multiproduct seller
Doyle (1986)	Yes	No	No	No	No
Narasimhan (1988b)	Yes	No	Yes with fixed promotion cost	No	No
Freimer and Horsky (2003)	Yes	No	No	No	No
Bergemann and Välimäki (2006)	No	Yes	Probability of low price	No	No
This paper	Yes	Yes	Yes	Yes	Yes

equivalent interpretation to learning and forgetting, as discussed above). Bergemann and Välimäki (2006) show 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). The fact that the low price is charged with some probability can be interpreted as a sale to attract the uninformed consumers. This paper concentrates on case (1) above, the uninformed consumers expected valuation is low compared to the optimal monopoly price to the informed consumers, and can be seen as applying the Bergemann and Välimäki (2006) type of analysis to gain insights on price promotions. In relation to Bergemann and Välimäki (2006), the paper first considers a simple two-period overlapping generations model to derive the result that temporary sales are offered (price cycles). Second, the paper considers a longer time horizon in discrete time with immediate learning 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 Bergemann and Välimäki 2006) and adds menus 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, because sales are typically offered for some length of time (e.g., several promotions last up to four weeks or longer), and not only for the minimum time period needed to cut and raise prices (which could be seen as less than a day or a week). Fourth, the analysis extends itself to obtain results for the case in which the seller carries more than one product (Villas-Boas and Villas-Boas 2006).

Three other papers are closely related to the results presented here: Doyle (1986), Narasimhan (1988b), and Freimer and Horsky (2003). Doyle (1986) 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 is not necessary for the existence of sales when consumers learn about their preferences for the products through experience. Narasimhan (1988b) 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. This can be seen as a trial-repeat model where price promotions are used to induce trial. In relation to Narasimhan, this paper formalizes the consumer preferences, allowing consumers to be strategic (which has implications on sale prices, and time interval between sales); studies the issue of the optimal duration of a sale; and can be easily extended to consider the multiproduct firm. Furthermore, the paper formalizes what is learned through trial, and how low prices can induce purchases given the consumer expectations over what can be learned, and explains the higher prices after consumers become informed. Freimer and Horsky (2003) 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. In contrast, this paper formalizes the consumer preferences and the learning behavior, and studies the optimal interval between sales, and the duration of a sale. The paper also shows that competition is not essential for the existence of sales with consumer learning. Table 1 summarizes this discussion highlighting the relation of the paper with this literature.

The remainder of this paper is organized as follows. Section 2 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; §4 considers the issue of the optimal length of a sale. Finally, §5 concludes. The proofs are in the technical appendix, which is provided in the e-companion.[6]

---

[6] An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

2. A Simple Model of Sales with Learning and Forgetting

This section presents a simple model illustrating how consumer learning and forgetting may lead to the existence of sales in a market. Consider a monopolist selling a nondurable, nonstorable 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 α of consumers derives utility u of consuming the product, whereas 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 the 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.[7] 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, whereas the price to attract the uninformed consumers is strictly less than u. This is the case in Bergemann and Välimäki (2006), where the uninformed consumers expected valuation, αu, is low compared to the optimal monopoly price to the informed consumers, u.[8] 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 αu, and, therefore, the optimal price when attracting the uninformed consumers is αu. The equilibrium would then be with alternating prices between the regular price u and the sale price αu.

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 had 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 αu, it gets a net present value of profits equal to $\alpha u + \delta V(1)$, whereas if it charges any price strictly above αu, no consumer buys this period and the firm gets a net present value of profits equal to $0 + \delta V(0)$, where δ is the discount factor. Charging a price of αu is better if $\alpha u + \delta[V(1) - V(0)] \geq 0$.

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 αu, 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 αu if $x = 0$, we have $V(0) = \alpha u((1 + \delta/2)/(1 - \delta^2))$ and $V(1) = \alpha u((1/2 + \delta)/(1 - \delta^2))$. Checking the no-deviation conditions above, one can then obtain that no deviations are profitable from the candidate equilibrium if $\alpha \leq \delta/(1 + \delta)$. That is, under this condition we get that the monopolist optimal behavior involves sales every two periods at the low price, αu. 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).

---

[7] 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.

[8] Note that in a two-point distribution of the consumer experiences we could still get the other case in Bergemann and Välimäki (2006), 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, $\alpha u + (1 - \alpha)\underline{u}$. Note also that in the current model with $\underline{u} = 0$, the firm is indifferent between all consumers being informed about their valuations, or all consumers being uninformed, because both yield the same profit of αu (as some of the informed consumers realize that they value the product at zero). If $\underline{u} > 0$, the profit of all consumers being uninformed would be greater than the profit of all consumers being informed, given that consumers are assumed to be risk neutral.

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), because there are few consumers to inform about the product. If consumers are in the market every period, one may argue that consumers are unlikely to forget about product fit, $1 - \beta$ is small. The results below follow along because the rate of forgetting is not zero. Alternatively, as suggested above, $1 - \beta$ could be seen as the rate at which consumers leave and come into the market, or the rate at which consumer preferences or product characteristics may change from period to period.[9] As discussed in the next section, not having immediate learning can be interpreted as consumers not being in the market in every period, which can potentially lead to more forgetting. Consumers can be seen as more likely to forget about the value of the product for categories with longer purchase cycles, for products for which consumers have less involvement, or for products with greater complexity.

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 restarts. 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 α 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 $\alpha u - p + \alpha(u - p)(\delta\beta)^T(1/(1 - (\delta\beta)^T)) \geq 0$, which results in

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

The sale price p that attracts the uninformed consumers is greater than the expected utility of consumption in a period, αu, because, once informed, the consumer has the potential of getting a positive surplus in future periods when the seller again offers 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 δ.

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 deciding 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 $p(\alpha x_t + 1 - x_t) + \delta V(\beta)$, because the fraction of informed consumers in period $t + 1$ will be β, 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.,

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

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 $V(\beta) = \alpha\beta u + \delta\alpha\beta^2 u + \delta^2\alpha\beta^3 u + \cdots + \delta^{T-2}\alpha\beta^{T-1}u + \delta^{T-1}p(\alpha\beta^T + 1 - \beta^T) + \delta^T V(\beta)$,

---

[9] One interesting issue not investigated here is that the more consumers buy a product, the less likely they may be to forget about product fit. In the analysis this would lead to more state variables (rather than just the fraction of informed consumers in the previous period), which would complicate the analysis as briefly discussed at the end of the next section.

from which one can obtain

$$V(\beta) = \frac{1}{1-\delta^T}$$

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

Putting together (1)–(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.

PROPOSITION 1. *The market equilibrium involves sales every T periods, where T is obtained from (1)–(3). For the discount factor δ close to one and β large, the time interval between sales T is greater, the lower the rate of forgetting $1-\beta$, the lower the probability of product fit α, and the lower the discount factor δ. Under the same conditions, the discount price p is decreasing in the rate of forgetting $1-\beta$, and increasing in the probability of product fit α and in the discount factor δ.*

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. Furthermore, a lower rate of forgetting leads to a higher benefit of being informed, which yields a smaller discount needed for the sale price. 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, and a reduced need to cut the discount price. A greater discount factor yields fewer relative benefits of offering a sale because the present value of the future profits are not too hurt by delaying the sale. Similarly, because the present value of the benefits of learning about the product is greater with a greater discount factor, the firm can then offer a smaller discount on the sale price.[10] 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 $\delta = 0.999$ (per week), a rate of forgetting of 1% per week (a yearly forgetting rate of 40%), leading to $\beta = 0.99$, and a probability of product fit $\alpha = 10\%$. This would then lead to an equilibrium number of weeks between sales of nine weeks. The pattern of evolution through time of prices and unit sales generated by this model seems to fit some product categories well, generating, in addition, a pattern of evolution for the fraction of informed consumers.

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 would result from the maximization of (3), while taking into account that p is a function of T. Because 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 noncommitment interval between sales. This result is intuitive: in the noncommitment 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 a 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 Bergemann and Välimäki (2006) result in that the probability of offering a sale is greater under commitment than under noncommitment. 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 Bergemann and Välimäki (2006), 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 the price back again. However, in the real world, sales last longer than the minimum amount of time needed for two successive price changes. 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 $dx/dt = \dot{x} = -(1-\beta)x$, where $\beta < 1$ (and β can

---

[10] The comparative statics presented are for the region of the parameter values stated in the proposition. We numerically checked other parameter values and obtained the same comparative statics results.

be negative).[11] 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 λ represents the learning parameter. When λ 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 Bergemann and Välimäki (2006). 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 with the product), or because consumers do not always learn about their fit when they try the product.

In this model, 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, or instantaneous price promotions and regular prices), because 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 Bergemann and Välimäki (2006). 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 αu, because the probability of them learning their valuation approaches zero. This means that the sale price has to be αu.[12] Let $k/2$

---

[11] The parameter β can be negative in this continuous-time specification because $(1 - \beta)$ is the continuous-time rate at which consumers forget, which can be greater than one. For a time interval T, the fraction of consumers who do not forget is $e^{-(1-\beta)T}$. For example, if time is measured in years, a yearly forgetting rate of 40% is represented by $\beta = 1 + \log 0.6 = 0.49$, and a yearly forgetting rate of 65% has a negative β.

[12] Note that the sale price was above αu in the previous section because consumers were sure to learn their fit after the first experience. Note also that this also results in lack of seller commitment not being a problem in this case, which simplifies the analysis. Note that the commitment problem would recur if consumers always learned their fit in their first experience.

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 αu. 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

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

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 $x = (x_h e^{-(1-\beta)zT} - \lambda/(1 + \lambda - \beta))e^{-(1+\lambda-\beta)(t-zT)} + \lambda/(1 + \lambda - \beta)$ 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

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

from which one can obtain for z and β close to one that x_h is decreasing in T, z, and increasing in λ and β. 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, β, and λ 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. *Suppose k is small. The optimal duration of the cycle T is increasing in the menu costs k, and when $k \to 0$, the optimal duration of the cycle T converges to zero, but more slowly than k. When $k \to 0$, the fraction z of the duration of the cycle with the high price converges to $(1/\lambda)(2 - \beta - \sqrt{(1 - \beta)^2 + (1 - \beta)(1 + \lambda - \beta)/(1 - \alpha)})$. 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 λ, and the probability of good fit α. For the optimal z close to one, the optimal duration of the cycle T is decreasing in the forgetting rate $(1-\beta)$, the learning rate λ, the probability of good fit α, 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," or instantaneous price promotions and regular prices, as argued above (if $z < 1$, the interesting case). 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, because 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 Bergemann and Välimäki (2006) for the probability of offering a sale.

Consider now the comparative statics on the optimal duration of the cycle.[13] 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 λ, and the probability of good fit α, 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 time, 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, because 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 relatively more important for the firm to extract the surplus of the current informed consumers than to invest for the future, 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. 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 $\beta = 0.9$ (with time measured in years, resulting in about 1.7% rate of forgetting per week),[14] a maximum percentage of informed consumers if the low price is charged forever of $\lambda/(1+\lambda-\beta) = 96\%$, a probability of good fit of $\alpha = 10\%$, menu costs of changing prices of \$0.52 per price change (resulting in $k = \$1.04$), and 0.7% of the maximum potential revenues under three sales per year (resulting in $3k = 0.007\alpha u(\lambda/(1+\lambda-\beta))$).[15] This then results in optimal cycles of about 117 days, and optimal sales durations of about 20 days. For each set of parameters, one can also obtain the evolution through time of price, of unit sales, and of the fraction of informed consumers. To illustrate the role of λ, note that with $\lambda = 2$ and $\beta = 0.9$, starting with no consumers informed, one month of the sale price yields a 15% fraction of informed consumers. This value moves to 22% if $\lambda = 3$.

We 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 the consumer can learn

---

[13] Note that the optimal z is close to one if β is high enough, or if λ or α is low enough.

[14] As a reminder, the parameter β in this continuous-time model has a different meaning of β in the previous (discrete-time) section.

[15] These values of menu costs of \$0.52 per price change, and 0.7% of revenues are the estimates presented in Levy et al. (1997) for the supermarket industry.

not only through experience, but also through word of mouth of the informed consumers (in a nonlinear 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 nonlinearity 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.[16]

5. 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 the model can be extended to the case in which the seller carries more than one product with additional insights.

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 pricing to do an in-depth empirical test of the results 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.

6. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

Acknowledgments

References

Aghion, P., P. Bolton, C. Harris, B. Jullien. 1991. Optimal learning by experimentation. *Rev. Econom. Stud.* **58** 621–654.

[16] See, for example, Feinberg (2001), for a discussion of related issues in the context of dynamic advertising policies.

Anderson, E. T., D. I. Simester. 1998. The role of sale signs. *Marketing Sci.* **17** 139–155.

Anderson, E. T., N. Kumar, S. Rajiv. 2004. A comment on: "Revisiting dynamic duopoly with consumer switching costs." *J. Econom. Theory* **116** 177–186.

Bergemann, D., J. Välimäki. 2006. Monopoly pricing of new experience goods. *J. Political Econom.* **114** 713–743.

Blattberg, R. C., S. A. Neslin. 1990. *Sales Promotion—Concepts, Methods, and Strategies*. Prentice Hall, Englewood Cliffs, NJ.

Blattberg, R. C., G. D. Eppen, J. Lieberman. 1981. A theoretical and empirical investigation of price deals for consumer nondurables. *J. Marketing* **45** 116–129.

Conlisk, J., E. Gerstner, J. Sobel. 1984. Cyclic pricing by a durable goods monopolist. *Quart. J. Econom.* **93** 489–505.

Doyle, C. 1986. Intertemporal price discrimination, uncertainty and introductory offers. *Econom. J.* **96** 71–82.

Feinberg, F. 2001. On continuous-time optimal advertising under S-shaped response. *Management Sci.* **47** 1476–1487.

Freimer, M., D. Horsky. 2003. Try it you will like it—Does consumer learning lead to competitive price promotions? Working paper, University of Rochester, Rochester, NY.

Jeuland, A. P., C. Narasimhan. 1985. Dealing–temporal price cuts— By seller as a buyer discrimination mechanism. *J. Bus.* **58** 295–308.

Keon, J. 1980. The bargain value model and a comparison of managerial implications with the linear learning model. *Management Sci.* **26** 1117–1130.

Kotler, P., K. L. Keller. 2006. *Marketing Management*. Pearson Prentice Hall, Upper Saddle River, NJ.

Kuksov, D. 2004. Buyer search costs and endogenous product design. *Marketing Sci.* **23** 490–499.

Lal, R., J. M. Villas-Boas. 1998. Price promotions and trade deals with multi–product retailers. *Management Sci.* **44** 935–949.

Lazear, E. 1986. Retail pricing and clearance sales. *Amer. Econom. Rev.* **76** 14–32.

Levy, D., M. Bergen, S. Dutta, R. Venable. 1997. The magnitude of menu costs: Direct evidence from large U.S. supermarket chains. *Quart. J. Econom.* **112** 793–825.

Macé, S., S. A. Neslin. 2004. The determinants of pre- and postpromotion dips in sales of frequently purchased goods. *J. Marketing Res.* **41** 339–350.

Mahanjan, V., E. Muller. 1986. Advertising pulsing policies for generating awareness for new products. *Marketing Sci.* **5** 89–106.

Mehta, N., S. Rajiv, K. Srinivasan. 2004. Role of forgetting in memory-based choice decisions. *Quant. Marketing Econom.* **2** 107–140.

Nagle, T. 1987. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*. Prentice Hall, Englewood Cliffs, NJ.

Narasimhan, C. 1988a. Competitive promotional strategies. *J. Bus.* **61** 427–449.

Narasimhan, C. 1988b. A model of discounting for repeat sales. T. M. Devinney, ed. *Issues in Pricing—Theory and Research*. Lexington Books, Lexington, MA.

Padilla, A. J. 1995. Revisiting dynamic duopoly with consumer switching costs. *J. Econom. Theory* **67** 520–530.

Pashigian, B. P. 1988. Demand uncertainty and sales: A study of fashion and markdown pricing. *Amer. Econom. Rev.* **78** 936–953.

Raju, J. S., V. Srinivasan, R. Lal. 1990. The effects of brand loyalty of competitive price promotional strategies. *Management Sci.* **36** 276–304.

Rao, R. C. 1991. Pricing and promotions in asymmetric duopolies. *Marketing Sci.* **10** 131–144.

Rao, R. C., R. Arjunji, B. P. S. Murthi. 1995. Game theory and empirical generalizations concerning competitive promotions. *Marketing Sci.* **14** G89–G100.

Rosenthal, R. 1980. A model in which an increase in the number of sellers leads to a higher price. *Econometrica* **48** 1575–1579.

Shilony, Y. 1977. Mixed pricing in oligopoly. *J. Econom. Theory* **14** 373–388.

Silk, A. J., G. L. Urban. 1978. Pre-test market evaluation of new packaged goods: A model and measurement methodology. *J. Marketing Res.* **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. *Amer. Econom. Rev.* **70** 651–659.

Villas-Boas, J. M. 1993. Predicting advertising pulsing policies in an oligopoly: A model and empirical test. *Marketing Sci.* **12** 88–102.

Villas-Boas, J. M. 1995. Models of competitive price promotions: Some empirical evidence from the coffee and saltine crackers markets. *J. Econom. Management Strategy* **4** 85–107.

Villas-Boas, S. B., J. M. Villas-Boas. 2006. Learning, forgetting, and sales. Working paper, University of California, Berkeley.

Zhao, Y. 2006. Price dispersion in the grocery market. *J. Bus.* **79** 1175–1192.