Advertising and Consumer Awareness of a New Product*

Alicia Barroso†

Job Market Paper, November 2007

Abstract

The increase of a new product's sales is usually attributed to consumers becoming informed about the existence of the product. Advertising can accelerate this consumer awareness process. This paper evaluates this effect. I develop and estimate a structural model in which the consumer purchase decision is specified using a discrete choice model with variation in the choice set, according to the information diffusion of the new products. I also model the optimal price and advertising decisions of the firm, taking into account the dynamic effect of advertising on future sales (via an increase in the proportion of consumers aware of the product). The model is estimated using Spanish automobile data. The results suggest that advertising significantly enhances the information diffusion of new products and that firms take into account in their advertising decision this dynamics. The estimates show that advertising reduces the three year it takes for the information diffusion of a new product to half as long.

Key words: New Product, Advertising, Consumer Awareness Process, Consumer Choice Set, Discrete Choice Model, Structural dynamic model.

JEL Classification: M37, L13.

*I thank Dmitri Byzalov, Ulrich Doraszelski, Jordi Jaumandreu, Gerard Llobet, Ariel Pakes, the participants at the 34th Annual Conference of the EARIE, and the ESF/COST Workshop, as well as the seminar audience at Harvard University and CEMFI for helpful discussions . I gratefully acknowledge the hospitality of Harvard University and CEMFI support during the academic year 2006/07.

†CEMFI. C/Casado del Alisal 5, 28014, Spain. E-mail: abarroso@cemfi.es

1 Motivation

The term product life is a description of what happens to product sales over time. According to Gort and Klepper (1982) the standard product life cycle consists of four stages: introduction, growth, maturity and decline[1]. During the first stages (the introduction and growth stages), sales increase as a result of consumers becoming informed about the existence of the new product. In the maturity stage, the proportion of consumers aware of the product is stable, and later, in the decline stage, sales decrease due to the entry of improved competing products. Advertising can accelerate the information diffusion process for a new product, by informing the consumers of product existence[2]. Notice that the effect of advertising on sales is dynamic, since advertising will affect the proportion of consumers aware of the product in the future, and therefore will affect future sales. Usually, firms advertise their products the most at the beginning of the life cycle. The dynamic effect of advertising on consumer awareness could explain this pattern.

Evidence of such common strategy is found in the Spanish automobile market. The estimated effect of age on market share confirms the existence of the product cycle in this market and shows that model sales typically increase in the first two years. This result suggests that the information diffusion of a new product in the Spanish automobile market takes place mostly in the first two years. The expenditure of advertising a product during this period is higher, even though sales are lower than in subsequent periods. This behavior is especially remarkable in the first years when the average expenditure on advertising is around 15.1% higher than in the fourth year, while the average sales of a model are approximately 24.6% lower. Therefore, products are heavily advertised during the period when their information diffusion takes place. These facts are consistent with two important ideas: first, the existence of a consumer awareness process for a new product, and second, an important role of advertising in the information diffusion for a new product.

The goal of this paper is to evaluate the role of advertising in the information diffusion of a new product. To address it I develop and estimate a structural discrete-choice model, in which the consumer is aware of only a subset of all products (her choice set), and her purchase decisions are only among the products in her choice set. The choice set evolves according to the information diffusion process, which is affected by the advertising expenditure. I also model the optimal price and advertising decisions of the firm, taking into account the dynamic effect of advertising on future sales (via an increase in the

[1] Several variations of the product life cycle model have been developed to handle the evolution of the product. Although the same basic stages are apparent in all, they differ as to the number and the names of the stages.

[2] For frequently purchased products, much of the learning is due to consumer's own experience with the product, however for durable goods like cars, there is much less learning from own consumption experience, and more learning from other sources like advertising.

proportion of consumers aware of the product). Firms face a dynamic optimization problem. Because It is computationally unfeasible to explicitly solve this optimization problem for estimation purposes, I use the technique proposed by Berry and Pakes (2007). This technique is based on estimating the optimality conditions for the dynamic controls (advertising in my case), and does not require an explicit solution for this optimization problem. I estimate the three equilibrium relationships (demand, price and advertising equations) using Generalized Method of Moments and simulation techniques. I use the algorithm proposed by Berry, Levinsohn and Pakes (1995), and also propose a way to simulate consumer choice sets that deals with the dimensionality problem that arises due to the high number of products[3] commonly observed in differentiated product markets. I apply this analysis to the Spanish automobile market.

The analysis is performed using a monthly panel data from January 1990 to December 2000 (132 months) , with "model" as the elementary unit of analysis (257 distinct models were sold in the market, offered by 33 multiproduct firms). There were 180 models that entered the market, and 93 exits with a mean age of around 8 years. The information gathered for each model includes price, new car registrations (sales), model age, brand, advertising expenditures, mechanical design and equipment characteristics.

The results suggest that advertising significantly enhances the information diffusion of new products and that the firms take into account in their advertising decision this dynamic. The estimates show that advertising reduces the three years it takes for the information diffusion of a new product to half as long.

The next subsection reviews the related literature.

1.1 Related Literature

The standard discrete choice models prevalent in IO literature assume that consumers are aware of all the products, and as a result those models only address variation in the choice sets across markets (in fact, it is an important source of identification in these models). However, there is some recent research in IO that focuses on other sources of variation to estimate more realistic demand specifications. Anupindi, Dada, and Gupta (1998) and Conlon and Mortimer (2007) study variation in consumer choice sets generated by the presence of stockouts. In Katz (2007), the variation comes from the fact that consumers restrict their attention to a subset of products before making a choice. This last paper is close to a large body of literature in marketing known as consideration set literature, focused on incorporating the variation in the consumer choice set into discrete choice models (Manski 1977 was the first to introduce it). In this literature, two interpretations of the choice set are possible. First, consumers might be unaware of the existence of some products, and their choice set consists of all the products they are aware of. Alternatively, consumers might face cognitive costs or constraints of having to consider a large number of

[3] The number of possible choice sets increases exponentially with the number of products.

products in their choice, and therefore they might restrict their attention to a smaller subset of products before making a choice. Both interpretations have been considered in the literature to study the effect of advertising on the consumer choice set. Goeree (2007) develops and estimates a model to analyze the effect of advertising on the consumer choice set for the PC market in the United States, where the consumer choice set is interpreted as all the products the consumer is aware of. Meanwhile, Chiang, Chib and Narasimhan (1999) assume the second interpretation and propose an integrated brand choice model that is capable of accounting for the heterogeneity in the choice set and in the parameters of the brand choice model. Both papers focus on the variation in the choice set across consumers according to consumer characteristics, and employ a static model.

My research considers the first interpretation of the choice set (all products a consumer is aware of), and the variation in the consumer choice set comes from the knowledge of the existence of new products. This variation, unlike in previous research, introduces dynamics in demand, which means a long-run effect of advertising through the consumer choice set.

One of the main problems of introducing variation in the consumer choice set (other than across markets) is the dimensionality problem that arises from the high number of possible choice sets. There are different strategies in the literature to deal with this problem, depending on the application and the data. For example, in Conlon and Mortimer (2007), the dimensionality problem arises from the fact that the exact timing of product stockouts (and therefore the choice set facing the consumer) is not observed. This feature of the data leads them to use the EM algorithm of Dempster, Lair, and Rubin (1997). Chiang, Chib and Narasimhan (1999) propose an estimation procedure using an approximation-free Markov chain Monte Carlo procedure.

This paper proposes a way to simulate consumer choice sets suited to deal with the dimensionality problem when the variation in the choice set comes from being unaware of the new products.

My paper is also related to the advertising literature. Two different effects of advertising on consumer choice have been analyzed[4]*:* the effect on consumer utility (Dixit and Norman, 1978) and the effect on the choice set (Grossman and Shapiro, 1984). Both of them have been previously studied in a static setting. In a dynamic setting, the effect of advertising on the utility has been considered theoretically (Friedman 1983, Chintangunta 1993, or Cellini and Lambertini 2003) and empirically (Dube, Hicks and Manchanda 2005, Ackerberg 2003 or Tan 2007). Most of them follow Nerlove and Arrow (1962), taking goodwill (the effect of advertising on the utility) to be a function of the stock of current and past advertising.

The dynamic effect of advertising through the consumer choice set, however, has never been considered in an empirical paper. In fact, Doraszelski and Markovich (2007) are the first to consider, in a theoretical

[4]In the literature, there is also a distinction made between the "informative" effect and the "persuasive" (or "prestige") effect. My reseach focuses on a different effect of advertising. I distinguish between the informative effect of advertising on the choice set, when a product enters the market, and the effect of advertising on utility.

framework, the long-run effect of advertising through the consumer choice set (which they call "awareness advertising"). They propose two different dynamic models of advertising, the goodwill and awareness advertising models, to study the effect of advertising on industry structure.

My research contributes to this literature by quantifying the long-run effect of advertising through the choice set when new products enter the market.

Only a few papers have directly addressed the life cycle of products introduced by multiproduct firms in a differentiated product market. Moral and Jaumandreu (2007) is the closest to this paper in this respect. Working within a discrete choice demand framework they introduce time effect as a product characteristic to study the demand price elasticities over product live cycle in the Spanish automobile market. Another example for an industry with a high product turnover due to intense innovation is Bresnahan, Stern and Trajtenberg (1997). They study the impact of being a frontier (technological) brand in the demand price elasticities. In my paper, unlike these papers, the information diffusion over the product cycle is explicitly modeled by introducing variation in the consumer choice set in the same demand framework.

This paper contributes to the literature in the following ways: (i) introducing variation in the consumer choice sets that evolve over time according to the information diffusion of new products, (ii) proposing a method to deal with the dimensionality problem that arises due to the high number of possible choice sets, (iii) evaluating the effect of advertising over the product life, taking into account the effect of advertising on the future proportion of consumers aware of the product, and (iv) being the first paper that applies the Berry and Pakes technique (2007) to real data.

The rest of the paper is organized as follows. I start by describing the data in Section 2. Sections 3 and 4 present the model and its implications (the consumer's and firm's problems, respectively); Section 5 discusses estimation issues; Section 6 presents the results and examines implications. Section 7 concludes.

2 Data

The model is estimated for Spanish car market data. It is an unbalanced panel data on a monthly basis from January 1990 to December 2000 (132 months), with "model" as the elementary unit of analysis. The original sources of the data set are ANFAC (Asociación Nacional de Fabricantes de Automóviles y Camiones), "Guía del comprador de coches" magazine and "Infoadex"[5].

In the data for every model, the most representative (sold) version is considered, and the characteristics of the model are taken from it. This rule leads to 257 distinct models offered by 32 brands (multiproduct

[5] Infoadex is a firm which computes advertising expenditure, by monitoring communication markets and their prices on a daily basis.

firms). Finally, treating a model/month as an observation, the total sample size is 16,362 observations. The information gathered for each model includes price, new car registrations (sales), segment, brand, advertising expenditures, mechanical design and equipment characteristics.

Infoadex reports total advertising expenditures on all models and automobile brands through the main media channels: newspapers, magazines, television, radio, cinema and billboards. In the automobile industry, it is common for firms to use both product and brand advertising. The (yearly) average expenditure on brand advertising is 18.2% of the total expenditure on advertising. Firms also use a combination of model-specific and group advertising, with groups of varying sizes. I divide group advertising expenditures by the number of models which compose the group, and construct model advertising expenditures by adding model-specific expenditures to these weighted group expenditures.

Other sources of information are INE[6] and EPA[7]. From them I collect the number of Spanish households (potential market size[8]) and the distribution of Spanish income per capita (annual mean and standard deviation[9]).

An interesting feature of the data is its high frequency, which helps to overcome the *data-interval-bias*. Clarke (1976) finds that estimated effects of advertising are sensitive to the frequency of data used, and calls it the data-interval-bias problem. The use of low-frequency advertising data when the effect of advertising on sales changes over a shorter period of time can lead to biased estimates of advertising effects.

The automobile industry is the second most important industry in the Spanish economy (after construction), representing almost 5 % of the GDP. After overcoming the second energy crisis, the eighties was a period of expansion for the Spanish car industry. Nevertheless, in the nineties, two environmental changes altered the evolution of the market: full integration into the EEC market, which implied a gradual reduction of tariff and non-tariff protection of the market (see Table 4), and the economic recession at the beginning of the decade (see Table 5).

Table 6 summarizes the evolution of the Spanish car market over the sample period. The market evolution during the nineties is characterized by (i) an increase in the number of car models, with a high rate of model introduction and turnover, (ii) an evolution of model sales closely related to the economic cycle, (iii) an important variation in the advertising expenditure, and (iv) stable prices, although model

[6] Instituto Nacional de Estadistica

[7] Encuesta de Población Activa

[8] The model market share is computed as unit sales of each model divided by the total market size (no. of households). Shares are annualized multiplying by 12 to facilitate comparability with the elasticities obtained with yearly data.

[9] The income distribution is assumed to be lognormal and its parameters are estimated from INE data. In particular, the estimated standard deviation $\widehat{\sigma}_y$ is 0.7 and the mean m_t, is the sample mean for each year.

characteristics continuously improved (the real price for a car with the same characteristics fell by 19% in the 90s). The model attributes considered are: size (m^2), auto cubic capacity per kg of model (cm^3/kg), gas mileage (kms covered at a constant speed of 90 kph with a liter of gasoline), and maximum speed (as measures of size and safety, power, fuel efficiency and luxury, respectively).

The entry of car models seems to be motivated by the intention to replace old models and introduce models absent until this time. During the nineties, 180 models entered the market and 93 exited, with a mean model age at exit of about 8 years.

The average annual advertising for a model is around 4 million euros, which means 4,65% of the revenue. Model advertising and model age have a close relationship. Table 1 indicates that models are intensely advertised for the first years, and similar conclusions are found with a simple OLS regression of model advertising on product attributes (Table 3). Some age effect also seems to be present for prices. However, it is much less economically significant in this market.

One interesting relationship is found between advertising and price. Table 7 shows the range of advertising over revenue ratio and the associated model, where the associated model is one whose price is the closest to the mean of the percentile. This relationship suggests a higher effect of advertising on the demand for the lower segments of the car market[10] (evidence addressed in Barroso, 2006).

3 The Consumer´s Problem

In the prevalent standard discrete choice model in IO, the assumption that the consumer is aware of all products implies that the market share of a product is the probability that it maximizes consumer's utility

$$\Pr\left(U_{ij} \geqslant U_{im},\ for\ m = 0, 1, ..., J | X_j, X_{-j}, Y_i\right)$$

where alternatives $m = 0, 1, ..., J$ represent purchases of the competing differentiated products. Alternative zero, or the outside alternative, represents the option of not purchasing any of those products.

When a model enters the market, however, certain proportion of consumers is not aware of the new product, and this proportion decreases over time with the diffusion of information. In this approach, the probability that product j will be purchased by consumer i is the probability that the consumer is aware of the product and that it maximizes her utility, U_i, given the subset of the J products that the consumer is aware of, S_i (consumer choice set)

$$\Pr\left(\{j \in S_i | X_j, X_{-j}, Y_i\} \cap \{U_{ij} \geqslant U_{im}, \forall m \in S_i\} | X_j, X_{-j}, Y_i\right)$$

[10] The simplest (static monopoly) Dorfman-Steiner condition stipulates that the advertising exposure ratio, a, over revenue (price p multiplied by quantity q) equals the advertising-demand elasticity ratio, η_a, over price-demand elasticity, η_p,

$$\frac{a}{pq} = \frac{\eta_a}{\eta_p} \Rightarrow \eta_a = \eta_p \frac{a}{pq}$$

where Y_i and X_j are consumer i and product j characteristics, respectively.

This probability can be written (see Appendix 1) as the sum of the probability that product j maximizes the utility given a choice set, multiplied by the probability of this choice set, for all the possible choice sets that include product j, Ω_j,

$$\sum_{S\in\Omega_j} \Pr\left(j \in S_i | X_j, X_{-j}, Y_i\right) \Pr(U_{ij} \geqslant U_{im}, \forall m \in S | X_j, X_{-j}, Y_i) \quad (1)$$

To get the probability that a product maximizes consumer's utility, the indirect utility for consumer i obtained from product j is assumed[11]

$$U_{ijt} = \delta_{jt} - \alpha_i p_{jt} + \epsilon_{ijt}, \text{ with } \delta_{jt} = \sum_k x_{jkt}\beta_k + \gamma a_j + \xi_{jt}$$

where x_{jt} is a K-dimensional vector of observed product attributes other than price p_{jt} and advertising expenditure[12] a_{jt}, the variable ξ_{jt} represents product characteristics unobserved (to the econometrician), and ϵ_{ijt} is a mean zero stochastic term which represents idiosyncratic individual preferences. As is customary in the literature, ϵ_{ijt} is assumed to be i.i.d. across products and consumers, and with the type I extreme value distribution. Consumer income is assumed to be distributed as $y_{it} \sim \log(m_t, \sigma_y^2)$ to model the price effect[13] as $\alpha_i = \alpha/y_i = \alpha e^{-(m_t + \sigma_y^2 v_{iy})}$, where the parameters m_t and σ_y^2 are the observed mean and variance, respectively (exogenous data).

I assume that the "outside option" (not buying any of the goods) is always included in the choice set. The utility from the "outside option", j=0, is

$$U_{i0t} = \sigma_0 v_{i0} + \epsilon_{i0t}$$

[11] This utility function comes from a Cobb-Douglas utility function in expenditures on other goods and services and characteristics of the good purchased:

$$U^*_{ijt} = (y_{it} - p_{jt})^\alpha G(x_{jt}, a_{jt}, \xi_{jt}) e^{\epsilon_{ijt}}$$

where $G(.)$ is assumed to be linear in logs, so that if $u_{ijt} = \log\left[U^*_{ijt}\right]$, then

$$\begin{aligned} u_{ijt} &= \alpha \log(y_{it} - p_{jt}) + \sum_k x_{jkt}\beta_k + \gamma a_j \delta_{jt} + \xi_{jt} + \epsilon_{ijt} \\ u_{i0t} &= \alpha \log(y_{it}) + \xi_{0t} + \sigma_0 \nu_{i0} + \epsilon_{i0t} \end{aligned}$$

If one normalize the $[\alpha \log(y_{it}) + \xi_{0t}]$ to zero, $U_{ijt} = u_{ijt} - [\alpha \log(y_{it}) + \xi_{0t}]$, then

$$U_{ijt} = (\alpha/y_{it}) p_{jt} + \sum_k x_{jkt}\beta_k + \gamma a_j \delta_{jt} + (\xi_{jt} - \xi_{0t}) + \epsilon$$

[12] There is a group of empirical studies that examine the impact of advertising on the price elasticity of demand, estimating interaction between price and advertising. Some examples are Krishnamurthi and Raj (1985) and Kanetkar, Weinberg and Weiss (1992). This research suggests that the effect of advertising on the price elasticity is difficult to determine and appears to vary across industries. However, this effect is usually associated with an advertising goodwill stock of brands and my research focuses on advertising of models where the long-run effect of advertising on the price elasticity intuitively seems less important.

[13] And so $v_{iy} \sim N(0, 1)$.

where $\sigma_0 v_{i0}$ is consumer i's taste for the outside good, and v_{i0} is an extra unobserved term that allows to account for the possibility that there is more unobserved variance in the idiosyncratic component of the outside than of the inside alternatives

Therefore, the conditional (on the consumer characteristics, $v_i = (v_{iy}, v_{i0})$, and the choice set, S) probability that product j maximizes consumer utility i is given by the standard logit expression

$$f_{jt|S}(v_i, \delta_t, X_t; \theta) = \frac{e^{\delta_{jt}+\alpha_i p_{jt}}}{e^{\sigma_0 v_{i0}} + \sum_{\forall m \in S} e^{\delta_{mt}+\alpha_i p_{mt}}} \tag{2}$$

To model the probability of a choice set, two assumptions are made. First, the probability that a product is included in the choice set does not depend on the consumer characteristics, and second it is also independent of the rival product attributes. So the probability of choice set S is the product of the probabilities for each product (see Appendix 1)

$$\Pr\left(j \in S_i | X_j\right) = \prod_{m \in S} \phi_{im} \prod_{n \notin S} \left(1 - \phi_{in}\right) \tag{3}$$

where the probability that product j is in the choice set in period t, ϕ_{jt}, is modeled as a concave function[14]

that depends on the level of awareness about the product,

$$\phi_{jt} = \frac{\omega_{jt}}{1 + \omega_{jt}} \tag{4}$$

The level of awareness about product j at time t, ω_{jt}, is assumed to exhibit time dependence, and it will be the previous level of awareness about the product, multiplied by the parameter φ, plus a random variable ς_{jt}, multiplied by the parameter ψ,

$$\omega_{j\ t+1} = \varphi \omega_{jt} + \psi \varsigma_{jt} \tag{5}$$

where the random variable $\varsigma_{jt} = 1$ with probability $\varkappa a_{jt} / \left(1 + \varkappa a_{jt}\right)$ and 0 otherwise.

The initial awareness level when a product enters the market ω_{0T_j} is assumed to be equal across products. Therefore, the parameters φ, $\varkappa$, ψ and the level of awareness when a model enters the market ω_0 will be parameters to be estimated.

This specification implies that the probability of being aware of a product is a function of the model age and the history of previous advertising expenditures (see Figure 5).

The predicted market share, or the awareness corrected market share, will be obtained by integrating out the expression (1) over the distribution of consumers' characteristics and the random variable of the awareness transition equation

[14] An alternative function that can be used is $\phi_{jt} = \frac{\exp(\omega_{jt})}{1+\exp(\omega_{jt})}$, a S-share function. It will have to be checked.

$$s_{jt}(P_0, \xi_t, X_t; \theta) = \int \sum_{S \in \Omega_j} \left(\prod_{m \in S} \phi_{mt} \prod_{n \notin S} (1 - \phi_{nt}) \right) \; f_{jt|S} \;\; dP_0(v_i, \varsigma_{jt})$$

where $\phi_{mt} = \phi(\varsigma_{jt}, X_t; \theta)$, $f_{jt|S} = f(v_i, \delta_t(\xi_t, X_t; \theta), X_t; \theta)$, and $dP_0(v_i, \varsigma_{jt})$ is the assumed distribution of (v_i, ς_{jt}). The consumer characteristics v_i are assumed to be independently normally distributed across the population and products with a mean of zero and a variance of one.

4 The Firm´s Problem

The firm f´s problem is to choose, for the set of models produced by $\Im_f$, a sequence of advertising expenditures, $\{a_{f\ t+\tau}\}$, and prices, $\{p_{f\ t+\tau}\}$, to maximize the expected discounted value of net cash flows, $\{\pi_{f\ t+\tau}\}$, conditional on product attributes, price, advertising and the levels of awareness of all (own and competing) product. So, the firm´s problem is to choose a_{ft} and p_{ft} to

$$sup_{a_{ft}, p_{ft}} E \left[\sum_{\tau=0}^{\infty} \rho^{\tau} \pi_{f\ t+\tau} | x_t, \omega_t, a_t, p_t \right]$$

where the expectation is taken with the understanding that optimal actions will be taken in each future period, ρ is the discount rate, and the net cash flow is given by

$$\pi_{f\ t} = \sum_{r \in \Im_f} \left[p_{rt} - mc_{rt} \right] s_{rt}(x_t, \omega_t, a_t, p_t) M_t - a_{rt}$$

where mc_{rt} is the marginal cost and M_t is the potential market size.

I assume firm f is a Bertrand competitor (in a differentiated-product market) that takes as given the characteristics of its products and competing products, as well as prices, advertising and the awareness levels. I also assume no asymmetric information. Then the vector $\{\omega_t\}$ will be observed by the firms[15] when the choices (advertising and prices) are to be made. This approach implies that price is a multiproduct static decision that satisfies the first-order condition

$$s_{jt}(x_t, \omega_t, a_t, p_t) + \sum_{r \in \Im_f} (p_{rt} - mc_{rt}) \frac{s_{rt}(x_t, \omega_t, a_t, p_t)}{\partial p_{jt}} = 0 \tag{6}$$

where the marginal cost is specified using the "hedonic" approach (take cost as a function of a set of product attributes; see, for example, Rosen 1974). In particular, I approximate the marginal cost using the log of product attributes, which are decomposed into a subset which is observed by the econometrician w_j, and an unobserved component ζ_j. Given these assumptions, the marginal cost of product j is written as

$$\ln(mc_j) = \ln(w_j)\eta + \zeta_j \tag{7}$$

[15] The realization of the awareness shock ς_{jt} is observed at the begining of the period t.

where η is a vector of parameters to be estimated.

Two additional assumptions are made to model the dynamic firm´s behavior on advertising. First, firm is only uncertain about the future awareness level of its product and competing products, although their distributions are known by all[16], and second, firm behaves as a single-product firm on advertising. The dynamic firm decision about advertising is modeled following a technique introduced by Berry and Pakes (2007), an alternative to the Euler equation techniques, in which it is not required to solve explicitly for the value function and the policy function of the firm so that implications of the firm´s choice can be used in estimation (see Appendix 2).

To apply Blackwell´s theorem, the standard assumptions that $0 \leq \rho < 1$, the state variable $\{\omega_t\}$ evolves as a Markov process conditional on the vector of action, and net cash flows are bounded from above are also made. Then the optimal advertising expenditure can be obtained from the unique value function that solves the Bellman equation

$$V_j(\omega_t) = sup_{a_{jt}} \left\{ \widetilde{\pi}_j(a_t, \omega_t) + \rho \sum_{\omega_{t+1}} V_j(\omega_{t+1}) dP(\omega_{t+1}|\omega_t, a_t) \right\}$$

where $\widetilde{\pi}_j(a_t, \omega_t) = \left[p^*_{jt} - mc_{jt}\right] s_j(\omega_t, a_t, p^*_t) M_t - a_{jt}$ is the indirect profit function given the equilibrium prices p^*_t (equilibrium relationship from the equation 6), with a_t and ω_t denoting the vector of advertising and level of awareness (respectively) of the firm itself and rival firms, and $P(\omega_{t+1}|\omega_t, a_t)$ is the Markov transition kernel for the vector[17] $\{\omega\}$.

If a_{jt} is an interior solution[18], it must satisfy the first-order condition:

$$0 = \frac{\partial \widetilde{\pi}_j(a_t, \omega_t)}{\partial a_{jt}} + \rho \sum_{\omega_{t+1}} V_{jt+1}(\omega_{t+1}) \frac{\partial p(\omega_{t+1}|\omega_t, a_t)}{\partial a_{jt}} d\omega_{t+1}$$

The dynamic term of the Bellman equation could be expressed as a expectation over a function of the value function in the next period

$$\sum_{\omega_{t+1}} V_{jt+1}(\omega_{t+1}) \frac{\partial p(\omega_{t+1}|\omega_t, a_t)}{\partial a_{jt}} d\omega_{t+1} = E\left[V_{jt+1}(\omega_{t+1}) \frac{\partial \ln p(\omega_{j\ t+1}|\omega_{jt}, a_{jt})}{\partial a_{jt}} |\omega_t\right]$$

Therefore the first-order condition for advertising could be rewritten as

[16] Note that future states and actions of rival models affect the firm-model value function since it is an oligopolistic market.

[17] Given the assumptions, the Markov transitions kerner for the levels of awareness are independent across products, so

$$\sum_{\omega_{t+1}} V_j(\omega_{t+1}) dP(\omega_{t+1}|\omega_t, a_t) = \sum_{\omega_{1\ t+1}} ... \sum_{\omega_{J\ t+1}} V_j(\omega_{1\ t+1}) dP(\omega_{1\ t+1}|\omega_{1t}, a_{1t}) ... dP(\omega_{J\ t+1}|\omega_{Jt}, a_{Jt})$$

[18] Only 5 models are not advertised: Renault 4, Jan1990-Aug1993, Skoda 136, Jan1990-Dec1991, Daewoo Compact, Jul1999-Dec2000, Toyota MR2, Mar2000-Dec2000, and BMW M3, Jan1997-Dec2000. 5.4% of the sample report zero advertising (usually no consecutives).

$$0 = \frac{\partial \widetilde{\pi}_j(a_t, \omega_t)}{\partial a_{jt}} + \sum_{\tau=1}^{\infty} \rho_{j\ t+\tau}^{\tau} \widetilde{\pi}_j(a_{t+\tau}, \omega_{t+\tau}) \frac{\partial \ln p(\omega_{j\ t+1}|\omega_{jt}, a_{jt})}{\partial a_{jt}} + \varepsilon_{jt} \tag{8}$$

$$\varepsilon_{jt} = \rho E\left[V_{jt+1}(\omega_{t+1}) \frac{\partial \ln p(\omega_{j\ t+1}|\omega_{jt}, a_{jt})}{\partial a_{jt}} |\omega_t\right] - \sum_{\tau=1}^{\infty} \rho^{\tau} \widetilde{\pi}_j(a_{t+\tau}, \omega_{t+\tau}) \frac{\partial \ln p(\omega_{j\ t+1}|\omega_{jt}, a_{jt})}{\partial a_{jt}} \tag{9}$$

where is the $p(\omega_{j\ t+1}|\omega_{jt}, a_{jt})$ is given by equation (5), and ε_{jt} is an expectations error.

The Rational Expectations assumption $E\left[\varepsilon_{jt}|\omega_t\right] = 0$ describes the equation that underlies the estimation[19]

$$E\left[\left((p_{jt} - mc_{jt}) \frac{s_{jt}(\omega_t, a_t, p_t)}{\partial a_{jt}} M_t - 1\right) + \sum_{\tau=1}^{\infty} \rho^{\tau} \widetilde{\pi}_j(a_{t+\tau}, \omega_{t+\tau}) \frac{1}{a_{jt}(1 + \varkappa a_{jt})} |\omega_t\right] = 0 \tag{10}$$

The equilibrium relationships (6) and (10) are a system of J equations since they are satisfied for all J products. To obtain the price and advertising equilibrium equations in matrix notation, I define two J by J new matrixes, the price effects matrix, Δ_t^p, and the brand matrix, Γ_t. The price effects matrix collects the own and cross-demand effects of price; therefore, its (r, j) element is $\Delta_{rj}^p = \frac{\partial s_{jt}}{\partial p_{rt}}$. The brand matrix is one of ones and zeros, with the (r, j) element being the indicator of whether product r is produced by the same firm as product j. I also define the vector Δ_t^a, which collects the own-demand effect of advertising where its j element (row) is $\Delta_{jt}^a = \frac{\partial s_{jt}}{\partial a_{jt}}$.

So, the equations (6) and (10) in matrix notation are

$$(p_t - mc_t) = \left[\Gamma_t \circ \Delta_t^p\right]^{-1} s_t \tag{11}$$

$$(p_t - mc_t) M_t \circ \Delta_t^a - I + \left(\sum_{\tau=1}^{T_j} \rho^{\tau} \pi_{t+\tau} \circ \frac{1}{a_t \circ (I + \varkappa a_t)}\right) = \varepsilon_t \tag{12}$$

where $\pi_{t+\tau} = (p_{t+\tau} - mc_{t+\tau}) M_{t+\tau} \circ s_{t+\tau} - a_{t+\tau}$, T_j the period when product j exits the market, the symbol $\circ$ represents Hadamard product, and I is a vector of J ones.

5 Estimation Strategy

The model consists of three equations, a demand, a price and an advertising equation, with the associated parameters $\alpha, \sigma_0, \beta, \gamma, \eta, \varkappa, \psi, \omega_0$ and φ (the discount factor, ρ, is fixed to be 0.99) to be estimated. They

[19] The level of awareness at the beginning of the period is given by $\omega_{j\ t+1} = \varphi\omega_{jt} + \psi\varsigma_{jt}$, where the random variable $\varsigma_{jt} = 1$ with probability $\varkappa a_{jt}/\left(1 + \varkappa a_{jt}\right)$ and 0 otherwise. So

$$\frac{\partial \ln p(\omega_{j\ t+1}|\omega_{jt}, a_{jt})}{\partial a_{jt}} = \frac{\partial \ln(\varkappa a_{jt}/\left(1 + \varkappa a_{jt}\right))}{\partial a_{jt}} = \frac{\partial(\ln(\varkappa a_{jt}) - \ln\left(1 + \varkappa a_{jt}\right))}{\partial a_{jt}} = \frac{1}{a_{jt}\left(1 + \varkappa a_{jt}\right)}$$

are estimated simultaneously by the generalized method of moments (GMM). I use moments arising from the demand and firms' pricing and advertising decisions, which express orthogonality between appropriate instruments and the unobservable components. Therefore, I consider the objective function $\Lambda' Z A_N^{-1} Z' \Lambda$, where A_N is a weighting matrix and $Z = (Z_\xi, Z_\zeta, Z_\varepsilon)$ are instruments orthogonal to the composite error $\Lambda = (\xi, \zeta, \varepsilon)$

$$Z'\Lambda = \begin{bmatrix} \sum_j Z_{\xi j}\xi_j(s^n, P_{ns}, P_R; \theta) \\ \sum_j Z_{\zeta j}\zeta_j(s^n, P_{ns}, P_R; \theta) \\ \sum_j Z_{\varepsilon j}\varepsilon_j(s^n, P_{ns}, P_R; \theta) \end{bmatrix}$$

where the observed vector of sampled market shares is s^n (number of households sampled n), the empirical distribution of ns simulation draws from the assumed distribution of consumer characteristics υ_i is P_{ns}, and P_R is the empirical distribution of R simulation draws from the assumed distribution of the state of awareness shocks ς_{jt}.

To compute the objective function, the unobservable components are solved following the technique proposed in BLP (1995). Given the number of models on the market, there are many possible choice sets, and a dimensionality problem arises. To deal with this, a simulation of consumer choice sets is introduced in the standard BLP algorithm. So, for each θ, I solve for unobservable components following the steps:

(i) Estimate, via simulation, the market shares implied by the model.

  (a) Given a draw from the Beurnoulli[20] ς_{jt} with probability $(\frac{\varkappa a_{jt}}{1+\varkappa a_{jt}})$, compute the vector of awareness probabilities for all products[21] $\phi_{jt}(\theta)$ according to the equation (4) and (5).

  (b) For each period t, simulate ns consumer choice sets as follows

    Define a Bernoulli[22] for each product, b_{jt}, with mean $\phi_{jt}(\theta)$

    Draw ns random variables from each Bernoulli, $(b_{jt1}, ., b_{jti}, ., b_{jtns})$

[20] The draws from the Bernouillis ς_{jt} are computed in the following way

$$\varsigma_{jtr} = \left\{ \begin{array}{lll} 1 & \text{if} & \varkappa a_{jt}/(1+\varkappa a_{jt}) \geq u_{jr} \\ 0 & \text{if} & \varkappa a_{jt}/(1+\varkappa a_{jt}) < u_{jr} \end{array} \right\}$$

where u_{jr} is the r draw from a uniform random variable for product j

[21] To compute the pre-sample awareness probability for models already existing at the beginning of the sample, I use the formula:

$$\omega_{j1/1990} = \omega_0 \varphi^{Age_J} + \psi \frac{\varkappa \overline{A}}{1+\varkappa \overline{A}} \sum_{s=0}^{Age_j - 1} \varphi^s$$

where $\overline{A}$ is the average monthly advertising for the model in 1990 .

[22] The draws from the Bernouillis b_{jt} are computed in the following way

$$b_{jti} = \left\{ \begin{array}{lll} 1 & \text{if} & \phi_{jt}(\theta) \geq u_{ji} \\ 0 & \text{if} & \phi_{jt}(\theta) < u_{ji} \end{array} \right\}$$

where u_{ji} is the i draw from a uniform random variable for the product j

Define the ns choice sets taking the i element from each draw, $S_{it} = (b_{1ti}, ., b_{jti}, ., b_{Jti})$

(c) Draw ns vectors (v_{iy}, v_{i0}) from a multivariate normal distribution with mean 0 and an identity covariance matrix

(d) Summarizing, for each period t, ns consumers are simulated, $(v_{iy}, v_{i0}, b_{1ti}, ., b_{jti}, ., b_{Jti})$, with the same characteristics but different choice sets over time.

This is repeated for H draws of ς_{jt} to compute the simulation estimator of the market shares

$$s_{jt}(P_{ns}, P_R, \delta_t, X_t; \theta) = \frac{1}{H} \sum_{h=1}^{H} \left(\frac{1}{ns} \sum_{i=1}^{ns} f_{jt|S_{it}}(v_i, \varsigma_{th}, \delta_t, X_t; \theta) \right) \tag{13}$$

(ii) Solve for the demand unobservables, implied by the simulated and observed market share. The demand unobservable is computed as

$$\xi_j(s^n, P_{ns}, P_H; \theta) = \delta_j(s^n, P_{ns}, P_H; \theta) - \left[\sum_k x_{jkt}\beta_k + \gamma a_{jt} \right] \tag{14}$$

where I solve for the mean utility and advertising effect $\delta_j(s^n, P_{ns}, P_H; \theta)$ recursively, using the contraction mapping suggested by BLP (1995), which matches the model-predicted market share and the observed market share s^n.

(iii) Calculate the vector of cost unobservables from the difference between price and the markup computed from the shares. Using the equation (7), the cost unobservables can be written as

$$\zeta(s^n, P_{ns}, P_H; \theta) = [p - b(s^n, P_{ns}, P_H; \theta)] - \ln(w)\eta \tag{15}$$

where $b(s^n, P_{ns}, P_H; \theta)$ is the markup calculated[23] from the equation (11) as

$$[\Omega \circ \Delta^p(s^n, P_{ns}, P_H; \theta)]^{-1} s^n$$

.

(iv) Calculate the rational expectations disturbance ε from the equation (12)

$$\varepsilon(s^n, P_{ns}, P_H; \theta) = (p_t - mc_t)M_t \circ \Delta_t^a - I + \left(\sum_{\tau=1}^{\infty} \rho^\tau \pi_{t+\tau} \circ \frac{1}{a_t \circ (I + \varkappa a_t)} \right) \tag{16}$$

[23] Note that in this approach the estimates price effect is
$\frac{\partial s_{jt}(P_{ns},P_R,\delta_t,X_t;\theta)}{\partial p_{jt}} = \frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{ns}\sum_{i=1}^{ns}\alpha e^{-(m_t+\sigma_y^2 v_{iy})}\ f_{jt|S_{it}}(v_i,\varsigma_{th},\delta_t,X_t;\theta)\left(1-\ f_{jt|S_{it}}(v_i,\varsigma_{th},\delta_t,X_t;\theta)\right)\right)$
$\frac{\partial s_{jt}(P_{ns},P_R,\delta_t,X_t;\theta)}{\partial p_{kt}} = \frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{ns}\sum_{i=1}^{ns}\alpha e^{-(m_t+\sigma_y^2 v_{iy})}\ f_{jt|S_{it}}(v_i,\varsigma_{th},\delta_t,X_t;\theta)\ f_{kt|S_{it}}(v_i,\varsigma_{th},\delta_t,X_t;\theta)\right)$
And the estimated advertising effect

$$\frac{\partial s_{jt}(P_{ns}, P_R, \delta_t, X_t; \theta)}{\partial a_{jt}} = \frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{ns}\sum_{i=1}^{ns}\gamma\ f_{jt|S_{it}}(v_i,\varsigma_{th},\delta_t,X_t;\theta)\left(1-\ f_{jt|S_{it}}(v_i,\varsigma_{th},\delta_t,X_t;\theta)\right)\right)$$

The method outlined requires instruments that are correlated with specific functions of the observed data, but are not correlated with the unobservables and the expectations disturbance. A common assumption made in the literature is that the supply and demand unobservables are mean independent of both observed product characteristics and cost shifters, $E\left[\xi_j \backslash (x, w)\right] = E\left[\zeta_j \backslash (x, w)\right] = 0$. Then the characteristics (including the awareness level) and cost shifter of products can be employed as instruments in this equation. Note first that price and advertising are not included in the conditioning vector (x, w). This is because price and advertising are determined in part by ξ and ζ, and therefore advertising and price will likely be correlated with the error term, making these variables endogenous. It is also important to realize that the vector (x, w) include*s* the characteristics and cost shifter of all products. One natural feature of oligopoly is that markups respond differently to own and rival products. Therefore, I distinguish the instruments between the characteristics of a product produced by the same multiproduct firm versus the characteristics of a product produced by rival firms. The time dimension of my data is short in relation to the variation space of product attributes, so the main source of the identification will be the cross-section. However, sometimes this is not enough, so I also use the following instruments: the differences in the prices with respect to their individual time means, $\widetilde{p_{jt}} = p_{jt} - (\frac{1}{T})\sum_{s} p_{js}$, lagged a number of periods. Instruments of this type were first proposed by Bhargava and Sargan (1983), and their moment restrictions have been studied in Arellano and Bover (1995). The error term in the advertising equation, and the equation from the firm's advertising decision which underlies the estimation is an expectation error. According to the rational expectation assumption it is not correlated with product attributes. Therefore I use product advertising and price as instruments in this equation.

To reduce computation time, I restrict the non-linear search over the parameters to a subset $\{\alpha, \sigma_0, \omega_0, \varkappa, \varphi, \psi\}$. I concentrate out the parameters $\{\beta, \eta\}$ and minimize the GMM objective function with regard to $\{\alpha, \sigma_0, \omega_0, \varkappa, \varphi, \psi\}$. This search is performed using the Nelder-Mead (1965) non derivative simplex search routine[24].

In order to obtain an optimal estimator, I use a consistent estimate of the weight matrix from a preliminary suboptimal GMM estimator (where $A = Z'Z$)

$$\widehat{A}_N = N\left(\widehat{\Omega} \otimes Z'Z\right)$$

where $\widehat{\Omega}$ is the preliminary suboptimal GMM residual covariance matrix. So, all the reported statistics are robust to heteroskedasticity and serial autocorrelation.

[24] The Nelder-Mead method is a commonly used nonlinear optimization algorithm which uses the concept of a simplex (a polytope of N+1 vertices in N dimensions). A non derivative method like this is required in this estimation given that the objective function is not smooth with respect to the awareness parameters (ω_0, $\varkappa, \psi$ and φ).

5.1 The Identification of the Effect of Advertising

Alternatively[25], the probability that product j will be purchased by consumer i (1) can be written as:

$$\Pr(U_{ij} \geqslant U_{im}, \forall m \in S_i | j \in S_i, Y_i, X_j, X_{-j}) \Pr(j \in S_i | Y_i, X_j)$$

where advertising is included as a product characteristic in X_j, and affects both probabilities (the advertising effect on consumer utility and consumer awareness about the product).

Such a structure of the choice probabilities implies that an increase in advertising of product j will increase both terms. However, at the end of the information diffusion process,when almost all consumers are aware of the product, only the first term will be affected. In this case, the only effect of advertising is through the consumer utility, because most consumers are already aware of the product.

5.2 Truncation Remainder

For models that are still alive at the end of the sampling period, $\sum_{\tau=1}^{\infty} \rho^{\tau} \pi(\omega_{t+\tau}, a_{t+\tau}; \theta)$ is not observed. That is, I have complete data on models which have exited. If one selects data only from those models which do sell off before the end of the sample, one would be selecting models on the basis of the sample random draws that determine the residual from the estimating equation, and hence incur the possibility of a selection bias[26]. Of course, the truncation remainder will affect the results depending on the sample, and a priory one is not concerned about correcting for the truncation remainder.

Therefore, the estimation strategy is to select data only from those models with complete data (models that exit before the last period of the sample) for the advertising equations (16). Later, however, we will need to see if the truncation remainder is a serious problem in the data set, and then correct for it if it is[27].

My data consists of an unbalanced panel data of 257 models, and only 97 models exit before the last period of the sample (Dec. 2000). It consists of 5,732 observations (out of the total of 16,362). The entire model live is observed for 45 of these models, which means they enter and exit during the sample period (there are 180 total model entries, so 135 new models stay on the market after the last period of the sample). The mean age of the model when it exits the market is around 8 years, although the heterogeneity is high (minimum age 2 years, and maximum age 20 years).

[25] See Appendix 1.

[26] Let T be the final year of the sample. The truncation remainder problem is that $V(\omega_{T+1})$ is not observed for those firms which survive until T, and the firms which survive until T are selected (in part) on the basis of the realizations of the disturbance terms.

[27] Berry and Pakes (2007) introduce a semiparametric technique which allows us to do that (checks and corrects it).

6 Empirical Results

In this section, the estimation procedure described is applied to the Spanish automobile market data set. First I begin by introducing different specifications (or scenarios) I will compare. The first scenario is where it is assumed that consumers are aware of all the products, and advertising only affects the consumer utility. This represents a standard assumption in the differentiated products literature, which I report for a random coefficients logit model (a little different from those typically estimated in the IO literature since a static advertising equation is included to deal with the endogeneity problem that arises from the effect of advertising on consumer utility). I will refer to this model as the standard model. In the second specification, I adjust the choice set for changes in the product awareness using the method outlined above, although without considering the effect of advertising on it (not including the random variable ς in the transition equation for the level of awareness), awareness probabilities only depend on product age. I will refer to this model as Age depended choice set model. Finally, the effect of advertising on the consumer choice set is added in the third specification, denoted as full model. Therefore, a dynamic effect of advertising arises and the firm's decision is estimated according to the equation (10)).

The result are presented as following: first, I discuss product differentiation and the substitution patterns present in the Spanish automobile market. Next, I discuss result which highling the importance of introducing the awareness process to evaluate the advertising effect over the product life. Finally, I discuss the effect of advertising on the welfare of both consumers and producers, and policy implications.

Table 8 reports the parameter estimates for the three models. The estimates are broken down into three categories. First, the demand side parameters that include the effect of product attributes on the mean consumer utility (β), the term on price (α), and the deviation of the consumer taste for the outside good (σ_0). Second, the effect of (ln) attributes (η) on the marginal cost of product are presented as the cost side parameters. Finally, the awareness parameters that underline the transition equation of the awareness level of products (which defines the awareness probabilities).

Product differentiation and the substitution patterns. Regardless to the model, the estimated coefficients of attributes (β) imply a significant positive effect on the mean utility. The coefficient of advertising (α) is significantly smaller in the age-dependet choice set model than in the standard model. This gap comes from the fact that if one does not control the information diffusion process, the initial growth of the sales will be attributed to the high level of advertising when products enter the market. Table 10 presents the demand-advertising elasticities attributed to the effect of the advertising on the utility, for the associated models to each percentiles of the ratio advertising over revenue. The associated model is one whose ratio is the closest to the mean of the percentile. These elasticities are estimates for both the standard model and the age-dependet choice set model. The (model) mean demand-advertising elasticity attributed to the effect of the advertising on the utility using the age-dependet choice set model is around 0.18, and it is overestimated 45% using the standard model.

The estimated effect of Price (α), implies a significant negative effect of price on the consumer utility. The price term, like the cost side parameters (η), do not change a lot across the specifications. According to the estimations, the (ln) attributes have a positive impact on the (ln) marginal cost of a product. All of them are significant, with the exception of the size that is highly correlated with the product weight. The demand-price elasticities estimated from these models reproduce the expected pattern in a differentiated-product market. I mean, higher mark-up for models with smaller demand-price elasticity. Both elasticity and mark up are presented in table 9 for the associated models to each percentiles of the price. The associated model is one whose price is the closest to the percentile. The reported elasticities are for the age-dependet choice set model. The (model) mean demand-price elasticity is 3.11. This results are according to previous literature for the car market (see for example Berry, Levinsohn, and Pakes 1995 or Jaumandreu and Moral 2007).

The effect of advertising on the awareness process. The awareness parameters for the age-dependet choice set model and the full model are presented in the last rows of Table 8. The initial awareness level (ω_0) does not change a lot between models. The effect of previous awareness level (φ), however, is smaller when the effect of the advertising in the awareness process is introduced. This result is the expected since in the first model the coefficient of the previous awareness level captures some of the advertising effect. To evaluate the awareness parameters of the advertising effect (ψ and $\varkappa$), Figure 6 represents, given the estimated parameters, the awareness probabilities over product age, set: first, no advertising exposure during the product live, and second, considering for each model age the (model) average expenditure on advertising observed in the data. The figure shows that without advertising the information diffusion of a new product will finish (most consumers know the product or the product awareness probability is close to one) around the third year. The observed expenditure on advertising reduces this process to the half. Therefore this result support the idea that advertising significantly reduces the information diffusion of new products.

To quantify the dynamic effect of advertising, I take the model Fiat Punto as an example. I compute the number of current and future consumers attributed to advertising using the formula

$$\sum_{s=t}^{\infty} s_{jt}(P_{ns}, P_R, \delta_t, X_t, ; \theta \backslash a_t^{ns}) - s_{jt}(P_{ns}, P_R, \delta_t, X_t, ; \theta \backslash a_t = 0)$$

where $s_{jt}(P_{ns}, P_R, \delta_t, X_t, ; \theta \backslash a_t^{ns})$ the simulator estimator of the market (equation 13),and a_t^{ns} is the observed advertising expenditure. According to the model, the new consumers attributed to the effect of advertising on the utility is given by

$$\frac{1}{H} \sum_{h=1}^{H} \left(\frac{1}{ns} \sum_{i=1}^{ns} \gamma \; f_{jt|S_{it}}(v_i, \varsigma_{th}, \delta_t, X_t; \theta | a_t^{ns}) \left(1 - \; f_{jt|S_{it}}(v_i, \varsigma_{th}, \delta_t, X_t; \theta | a_t^{ns})\right) \right) a_t^{ns}$$

where $f_{jt|S_{it}}(v_i, \varsigma_{th}, \delta_t, X_t; \theta | a_t^{ns})$ is given by equation (2)

Therefore, the new consumers attribute to the effect of advertising on both the consumer utility and the consumer awareness process could be distinguished. Table 11 and Figure 7 present the results. We can observe that the new consumers attributed to the effect of advertising on the awareness process is significant and decreasing over the product live.

7 Conclusions

This paper evaluates the role of advertising in the information diffusion of a new product. Using a structural model in which consumer purchase decision is specified using a discrete choice model with variation in the choice set and a dynamic model of advertising taking into account the dynamic effect of advertising on future sales, my paper finds that advertising significantly enhances the information diffusion of new products and that the firms take into account the dynamic effect of advertising.

This paper focusses on durable products and one of the most important variables that determine the product sales, namely advertising. One of the main contributions of this paper is to show how the information diffusion of a product may be incorporated into an empirical model of consumer behavior that reproduces aspects of the product life cycle. Extensions of this model could adress questions like: how consumer learning affects sales?, What are the implications over the optimal firm's decision of entry and exit?, or welfare gains from the informative effect of advertising.

References

[1] Ackerberg, D. (2003), "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination", *International Economic Review,* 44 (3), 1007-1040

[2] Anupindi R., M. Dada, and S. Gupa (1998), "Estimation of Consumer Demand with Stock-Out Based Substitution: An Application to Vending Machine Products", *Marketing Science*, 66, 278-308.

[3] Arellano M. and P. Bover. (1995), "Another look at instrumental-variable estimation of error components models", *Journal of Econometrics,* 68, 29-51.

[4] Barroso A. (2006), "Advertising and Demand: Some Empirical Evidence in the Spanish Automobile Market", *Working paper.*

[5] Berry S.T., J. Levinsohn,. and A. Pakes. (1995), "Automobile Prices in Market Equilibrium", *Econometrica,* 63, 841-890.

[6] Berry S.T and A. Pakes (2007), "Estimation from the Optimality Conditions for Dynamic Controls", *Working paper.*

[7] Bhargava A., and J.D. Sargan (1983), "Estimating dynamic random effects models from panel data covering short time periods", *Econometrica,* 51, 1635-1659.

[8] Bresnaham, Stern and Trajtemberg's (1997), "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the late 1980s", *Rand Journal of Economics*, 28, 17-44.

[9] Cellini R. and L. Chambertini (2003), "Advertising with Spillover Effects in a Differential Oligopoly Game with Differentiated Goods", *Central European Journal of Operations Research*, 11, 409-423.

[10] Chiang J., S. Chib, and C. Narasimhan (1999), "Markov chain Monte Carlo and models of consideration set and parameter heterogeneity", *Journal of Econometrics*, 89, 223-248.

[11] Chintangunta P. (1993), "Investigating the Sensitivity of Equilibrium Profits to Advertising Dynamics and Competitive Effects", *Management Science*, 39 (9), 1145-1162.

[12] Clarke D. (1976), "Econometric Measurement of the Duration of Advertising Effect on Sales", *Journal of Marketing Research*, 13 (4), 345-357.

[13] Conlon C. and J. Mortimer (2007), "Demand Estimation Under Incomplete Product Availability", *Working paper.*

[14] Dixit A. and T. Wilson (1978), "Advertising and Market Power", *Bell Journal of economics*, 9 (1), 1-17.

[15] Doraszelski U. and S. Markovich (2007), "Advertising Dynamics and Competitive Advantage", *forthcoming in Rand Journal of Economics.*

[16] Dube J., G. Hitsch, and P. Manchanda (2005), "An Empirical Model of Advertising Dynamic", *Quantitative Marketing and Economics*, 3 (2), 117-144.

[17] Friedman J. (1983), "Advertising and Oligopolistic equilibrium", *Bell Journal of Economics,* 14 (2), 464-473.

[18] Grossman G. and C. Shapiro (1984), "Informative Advertising with differentiated products", *Review of Economic Studies*, 51 (1), 63-81.

[19] Goeree M. (2007). "Advertising in the US Personal Computer Industry", *Working paper.*

[20] Gort M. and S. Klepper (1982), "Time paths in the diffusion of product innovations", *Economic Journal,* 92, 630-653.

[21] Hansen L. and K. Singleton (1982), "Generalized instrumental variable estimation of nonlineal rational expectation models", *Econometrica*, 50, 1269-1286

[22] Kanetkar V., C. Weinberg, and D. Weiss (1992), "Price Sensitivity and Television Advertising Exposures: Some Empirical Findings", *Marketing Science*, 11, 359-371.

[23] Michael Katz (2007),

[24] Moral. M and J. Jaumandreu (2007), "Automobile demand, model cycle and age effects", *Spanish Economic Review*, 9 (3), 193-218.

[25] Krishnamurthi L. and S. Raj (1985), "The Effect of Advertising on Consumer Price Sensitivity", *Journal of Marketing Research*, 22, 119-129.

[26] Kwoka J. (1996), "Altering the Product Life Cycle of Consumer Durables: The Case of Minivans", *Managerial and Decision Economics*, 17, 17-25.

[27] Manski C. (1977), "The Structure of Random Utility Models", *Theory and Decisions*, 8, 229-254.

[28] Nelder, J.A. and r. Mead (1965), "A Simplex Method for Function Minimization", *Computer Journal,* 7 (4), 308-313.

[29] Nerlove M. and K. Arrow (1964), "Optimal advertising policy under dynamic considerations", Economica, 29 (114), 129-142.

[30] Tan W. (2007), "A Dynamic Analysis of the U.S. *Working paper.*

Appendix 1. The choice probability.

Consumer i chooses among $J+1$ products, where $j=0$ denoting the outside alternative and each product j has observable attributes X_j. Consumer i has characteristics Y_i and a choice set S_i (subset of the J products that contains all the products the consumer is aware of, where the outside alternative is assumed to be always known by consumers). Product j is known by the consumer i (included in the choice set S_i) with probability

$$\Pr(j \in S_i | Y_i, X_j)$$

that may vary across individuals and across products, and it may depend on interactions of consumer and product characteristics (for example, young consumers may be more likely to include the products most advertised on the Internet). Note the assumption below that $\Pr(j \in S_i|Y_i, X_j)$ does not depend on rival product characteristics X_{-j}. This assumption might rule out some of the cognitive constraints stories behind the consumer choice set. For example, if the consumer only remembers the 3 most advertised products, this assumption will be violated.

The consumer chooses one of the products from his choice set, so the conditional probability of choosing j (for $j \in S_i$) is

$$\Pr(j|S_i, Y_i, X_j, X_{-j})$$

where X_{-j} denotes the characteristics of all the products except j.

The choice probability for product j can be computed by integrating out all possible choice sets that include product j, $\mathcal{F}_j$,

$$\Pr(j|Y_i, X_j, X_{-j}) = \sum_{S_i \in \mathcal{F}_j} \Pr(j|S_i, Y_i, X_j, X_{-j}) \Pr(S_i|Y_i, X_j, X_{-j})$$

where

$$\Pr(S_i|Y_i, X_j, X_{-j}) = \prod_{m \in S_i} \Pr(m \in S_i|Y_i, X_m) \prod_{n \in S_i} (1 - \Pr(n \in S_i|Y_i, X_n))$$

For the discussion of identification, it will be more convenient to rewrite the choice probability as

$$\Pr(j|Y_i, X_j, X_{-j}) = \Pr(j|j \in S_i, Y_i, X_j, X_{-j}) \Pr(j \in S_i|Y_i, X_j)$$

where

$$\Pr(j|j \in S_i, Y_i, X_j, X_{-j}) = \sum_{S_i \in \mathcal{F}_j} \Pr(j|S_i, Y_i, X_j, X_{-j}) \Pr(S_{i,-j}|Y_i, X_{-j})$$

with $S_{i,-j}$ denoting the elements of consumer choice set other than j, and

$$\Pr(S_{i,-j}|Y_i, X_{-j}) = \prod_{\substack{m \in S_i \\ m \neq j}} \Pr(m \in S_i|Y_i, X_m) \prod_{n \notin S_i} (1 - \Pr(n \in S_i|Y_i, X_n))$$

Appendix 2. Berry and Pakes (2007, working paper) Technique

Berry and Pakes (2007, working paper) introduce an estimation technique based on the first-order conditions (optimality conditions) for continuous controls for dynamic models. The estimation technique is driven by the assumption that agent perceptions of the distribution of the discounted value of future net cash flows are rational in the sense that they are, at least on average, consistent with the realizations of those net cash flows (an assumption used extensively in applied work, standard in the literature).

It is applied to empirical models of markets where current decisions will have independent impact on future states, as well as on current profits. Examples include models with: learning by doing, experience, network or addictive goods, and collusion (applies both to models of single agents choosing policies in games against nature, and to dynamic games). This techniques also allows one to use the information on investment choice (advertising, R&D or investment in traditional capital stock) to help in estimation.

In the context of the single agent models, the investment equations are typically analyzed using the more standard Euler equation estimation techniques introduced into economics by Hall (1979) and Hansen and Singleton (1989). The technique proposed by Berry and Pakes (2007, working paper) is an alternative to this, with computational properties similar to those of Euler equation techniques. That is, neither of them requires the researcher to solve explicitly for the value function and/or the investment policy of the firm in order to use the implications of the investment choices in estimation. This alternative method has both advantages and disadvantages relative to the Euler equation technique. As advantages, (i) it allows us to analyze the dynamic game, (ii) it uses a stochastic accumulation model, and (iii) it allows future actions to be at the corner of the choice set, and so on. On the other hand, to use the Berry and Pakes (2007, working paper) method, one does have to specify the transition probabilities of the states conditional on the controls. It is likely, however, that this method is less sensitive to slight misspecifications in timing assumptions than is the Euler equation because it is based on the returns for investment over the entire future, while Euler equation techniques are only concentrated on returns over the period until the compensating perturbations are made.

Standard first-order conditions for continuous controls depend upon the derivative of the expected discounted value of future net cash flows. This expectation is not directly observed. However, under the rational expectation assumption, this can be expressed as a derivative of the discounted value of future net cash flows plus an error that will be mean independent of all variables known at the time the expectation is made. Then the first-order conditions will be a conditional expectation of known primitives, and a standard method of moments estimation algorithm can be used. To show the method, I present the firm's problem in a general setting.

The firm's problem is to choose a sequence of actions, say $\{a_{t+\tau}\}$, to maximize the expected discounted value of net cash flows, $\{n_{t+\tau}\}$, conditional on the information sets, $\{\omega_{t+\tau}\}$, that will be available when those actions are to be taken. So, the firm´s problem is to choose a_t to

$$sup_{a_t} E\left[\sum_{\tau=0}^{\infty} \rho^{\tau} n_{t+\tau} | \omega_t, a_t\right] \equiv V(\omega_t)$$

where the expectation is assumed with the understanding that optimal actions will be taken in each future period, and ρ is the discount rate. To apply Blackwell´s theorem, we also make the standard assumptions that $0 \leq \rho < 1$, $\{\omega_t\}$ evolves as a Markov process conditional on the vector of action, and net cash flows are bounded from above. Then the optimal action can be obtained from the unique value function that solves the Bellman equation

$$V(\omega) = sup_a \left\{ n(a,\omega) + \rho \int_{\sigma'} V(\omega') dP(\omega' | \omega, a) \right\}$$

where $P(\omega'|\omega, a)$ is the Markov transition kernel for $\{\omega_t\}$ conditional to the action a.

Two assumptions are made:

- Rational Expectations:

$$\sum_{\tau=0}^{\infty} \rho^{\tau} n_{t+\tau} = V(\omega_t) + \varepsilon_t$$
$$\text{with, } E\left[\varepsilon_t | \omega_t\right] = 0$$

- Smoothness: (i) $P(\omega' \backslash \omega, a)$ has support which is independent of a, (ii) the density $p(\omega'|\omega, a)$ differentiable in a for almost every ω' and every ω, and (iii) $n(\omega, a)$ is a differentiable function of a for almost every ω.

Given the Bellman Equation and the assumptions:

$$\begin{aligned} 0 &= \frac{\partial n(\omega_t, a_t)}{\partial a_t} + \rho \int V(\omega_{t+1}) \frac{\partial p(\omega_{t+1}|\omega_t, a_t)}{\partial a_t} d\omega_{t+1} \\ &= \frac{\partial n(\omega_t, a_t)}{\partial a_t} + \rho E\left[V(\omega_{t+1}) \frac{\partial lnp(\omega_{t+1}|\omega_t, a_t)}{\partial a_t} | \omega_t\right] \end{aligned}$$

$$0 = \frac{\partial n(\omega_t, a_t)}{\partial a_t} + \sum_{\tau=1}^{\infty} \rho^{\tau} n_{t+\tau} \frac{\partial lnp(\omega_{t+1}|\omega_t, a_t)}{\partial a_t} + \varepsilon_t$$
$$\text{where } \varepsilon_t = \rho E\left[V(\omega_{t+1}) \frac{\partial lnp(\omega_{t+1}|\omega_t, a_t)}{\partial a_t} | \omega_t\right] - \sum_{\tau=1}^{\infty} \rho^{\tau} n_{t+\tau} \frac{\partial lnp(\omega_{t+1}|\omega_t, a_t)}{\partial a_t}$$

So the equation which underlies the estimation of the parameters is $E(\varepsilon_t|\omega_t) = 0$, that is

$$E\left[\frac{\partial n(\omega_t, a_t; \theta)}{\partial a_t} + \sum_{\tau=1}^{\infty} \rho^{\tau} n_{t+\tau}(\omega_{t+\tau}, a_{t+\tau}; \theta) \frac{\partial lnp(\omega_{t+1}|\omega_t, a_t; \theta)}{\partial a_t} | \omega_t\right] = 0$$

Table 1: IV Logit Demand

Dependent variable: $ln(s_j) - ln(s_0)$

Constant	-16.909	(0.377)	-17.352	(0.380)
CC/Weight	2.076	(0.123)	2.169	(0.116)
Max.Speed	1.274	(0.114)	1.390	(0.115)
Km/l	0.603	(0.075)	0.617	(0.073)
Size	2.767	(0.354)	3.061	(0.342)
Age			0.060	(0.071)
Age^2			-0.039	(0.017)
Age^3			0.005	(0.002)
Age^4			-2.18e-4	(0.84e-4)
Age^5			0.03e-4	(0.01e-4)
Price	-1.822	(0.080)	-1.869	(0.083)

Note: Seasonal, brand, and regulation plans dummies are included in the equation. Instruments: BLP instruments, differences of prices with respect to their (indiv) time mean lagged one year, and the number of model and new models. The standard errors (reported in parentheses) are robust to heteroskedasticity and serial correlation

Table 2: Model Age and the Expenditure on Advertising a Model (model means)

Interval of Model Age	Advertising (millions euros 1995)	Sales (units)	Advertising/revenue (%)
1 year	5.57	6,846.3	9.97
1-2 years	5.52	9,052.5	6.69
2-3 years	5.54	9,729.6	6.16
3-4 years	4.84	9,081.4	5.09
4-5 years	4.83	8,879.5	4.84
5-6 years	4.37	8,814.4	4.23

Note: To avoid model selection problems, only models that exit the market after six years are considered (72, 75, 79, 91, 93 and 94 models). The mean age of models that exit the market is around 8 years.

Table 3: Age Effect for Advertising and Price

	OLS ln(price) on ln(w)		OLS Adv on w	
Constant	3.723	(0.078)	-12.168	(2.875)
CC/Weight	0.518	(0.009)	1.234	(0.293)
Max.Speed	1.032	(0.014)	7.474	(4.282)
Km/l	0.243	(0.009)	3.934	(0.307)
Size	0.198	(0.017)	-0.306	(1.334)
Weight	1.055	(0.012)	-2.728	(0.524)
Trent	-0.001	(3e-5)	-1.345	(0.423)
Age	-0.020	(0.003)	-0.629	(0.183)
Age^2	0.006	(0.001)	0.066	(0.052)
Age^3	-7e-4	(1e-4)	-0.003	(0.005)
Age^4	0.3e-4	(0.05e-4)	1e-4	(3e-4)
Age^5	-0.02e-4	(0.01e-4)	-0.02e-4	(0.05e-4)
N. Models			0.106	(0.032)
N. New Models			-0.029	(0.026)

Note: The estimated decrease in advertising over the first years is around 20%, while for price it is 0.01%. The standard errors are reported in parentheses

Table 4: Tariff in the Spanish automobile market (%)

	1987	1988	1989	1990	1991	1992	From 1993 on
European foreign producers	28.4	22.9	17.4	12.8	8.2	3.6	0
Non-European foreign producers	34.3	24.5	23.6	19.3	15.08	12.8	10

Table 5: Spanish income per capita (euros 1995)

1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
7749,5	7863,1	7949,6	7824,8	7836,1	8234,7	8257,5	8450,4	8655,8	9093,1	9232,3

Table 6: Entries/Exits and Basic Statistics (yearly model means)

	No. of Models	Model Entries	Model Exits	Quantity (sales)	Price	Advertising	Size (m^2)	Max. Speed	Kms/l	cc/w
1990	97	20	2	10,979.33	11,870	3.27	6.61	171.7	5.29	1.62
1991	105	10	5	8,730.8	11,740	2.64	6.66	173.1	5.32	1.61
1992	116	16	10	9,686.7	11,286	3.89	6.71	174.9	5.34	1.64
1993	117	11	8	6,615.8	11,557	4.72	6.74	175.9	5.34	1.63
1994	122	13	12	8,272.9	11,469	4.91	6.69	174.3	5.39	1.69
1995	127	17	11	7,316.3	11,844	4.18	6.70	175.1	5.49	1.56
1996	134	18	11	7,660.2	11,941	4.19	6.76	176.4	5.44	1.51
1997	152	29	11	7,902.1	11,907	4.13	6.84	178.3	5.54	1.48
1998	160	19	11	8,642.2	11,918	3.97	6.91	180.3	5.76	1.46
1999	160	11	7	9,317.3	11,741	3.89	6.97	182.3	5.81	1.44
2000	169	16	5	9,351.3	11,783	3.44	7.01	183.8	5.85	1.43
Total		180	93							

Note: Price (euros 1995) and characteristics are sales weighted means. Advertising is in millions of 1995.

Table 7: The Range of Advertising/Revenue (obs with age>2)

Percentile	(%) Adv/revenue	Associated Model	Price	Model Adv
0	0	BMW M3	52,158.2	0
10	0.21	Chrysler G. Voyager	25,453.5	0.03
20	0.91	Mercedes 190	25,233,9	0.67
30	1.63	Audi 80	20,147.6	1.07
40	2.50	Volvo 940	19,931.7	1.65
50	3.23	Citroen Xantia	14,296.0	2.75
60	4.12	Ford Puma	13,702.9	3.92
70	5.29	Honda Concerto	12,552.8	4.18
80	6.72	Toyota Carina	13,842.1	5.43
90	9.35	Fiat Tipo	11,842.7	7.99
100	34.85	Suzuki Swift Sedan	9,879.7	34.86

Note: Associated model is one whose price is the closest to the mean of the percentile. Price (euros 1995) and advertising (in millions of euros 1995) are the annual means of the associated model. Only 5 models are not advertised (5.4% of the sample report zero investment).

Table 8: Estimated Parameters of the Demand, Pricing and Advertising Equations

		Standard Model		Age-Dependet Choice Set Model		Full Model
Demand Side Parameters						
Means (β)	Constant	-8.561	(0.113)	-8.455	(0.083)	-8.555
	CC/Weight	0.058	(0.019)	0.346	(0.070)	0.107
	Max.Speed	1.605	(0.017)	1.656	(0.064)	1.968
	Km/l	0.063	(0.028)	0.246	(0.049)	0.014
	Size	4.752	(0.048)	3.781	(0.067)	5.064
	Adv	1.692	(0.014)	1.083	(0.099)	1.742
Std. Dev of the outside good (σ_0)		3.112	(0.001)	3.055	(0.001)	3.361
Term on Price (α)		-28.67	(0.554)	-28.83	(1.320)	-30.76
Seasonal controls		yes		yes		yes
Brand controls		yes		yes		yes
Regulation plans controls		yes		yes		yes
Cost Side Parameters						
Attributes (η)	Constant	4.299	(0.441)	4.306	(0.576)	4.292
	ln(CC/Weight)	0.566	(0.019)	0.564	(0.039)	0.579
	ln(Max.Speed)	0.820	(0.103)	0.818	(0.110)	0.818
	ln(Km/l)	0.206	(0.033)	0.209	(0.060)	0.219
	ln(Size)	0.121	(0.097)	0.128	(0.131)	0.112
	ln(Weight)	0.988	(0.020)	0.988	(0.076)	0.983
	Trend	-0.003	(0.001)	-0.003	(0.001)	0.002
Brand controls		yes		yes		yes
Awareness Parameters						
Initial Awareness (ω_0)				0.3489	(0.001)	0.342
Previous Awareness (φ)				1.355	(0.002)	1.181
Advertising Parameter ($\varkappa$)						0.753
Awareness Random Vb Parameter (ψ)						4.778

Note: the instruments are BLP (1995) instruments and differences of prices with respect to their (indiv) time mean lagged one year, and the number of model and new models. The standard errors (reported in parentheses) are robust to heteroskedasticity and serial correlation

Table 9: Demand-Price elasticities

Percentile	Price	Associated model	Price Elast	Markup (%)
0	0.491	Marbella	4.910	15.0
10	0.729	Micra	4.143	18.9
20	0.890	Logo	3.113	29.3
30	1.089	Scoupe	3.296	25.8
40	1.212	Vento	3.124	29.5
50	1.374	Marea	2.812	33.7
60	1.636	Lybra	2.490	38.7
70	1.888	Swift	2.596	38.4
80	2.192	Galaxy	2.995	31.4
90	2.528	325	2.396	42.1
100	6.919	TT	3.220	32.2

Note: Estimated elasticities from the Age-Dependet Model parameters.
Associated model is one whose price is the closest to the mean of the percentile.

Table 10: Demand-Advertising Elasticities from the effect on Consumer Utility

Percentile	Adv/revenue (%)	Associated Model	Standard Model	Age-Dependet Choice Set Model
0	0	BMW M3	0.000	0.000
10	0.21	Chrysler G. Voyager	0.027	0.017
20	0.91	Mercedes 190	0.045	0.029
30	1.63	Audi 80	0.029	0.019
40	2.50	Volvo 940	0.188	0.120
50	3.23	Citroen Xantia	1.093	0.700
60	4.12	Ford Puma	1.639	1.049
70	5.29	Honda Concerto	0.215	0.138
80	6.72	Toyota Carina	0.301	0.193
90	9.35	Fiat Tipo	0.183	0.117
100	34.85	Suzuki Swift Sedan	0.975	0.624

Note: Observations for model age bigger than two years are only considered.

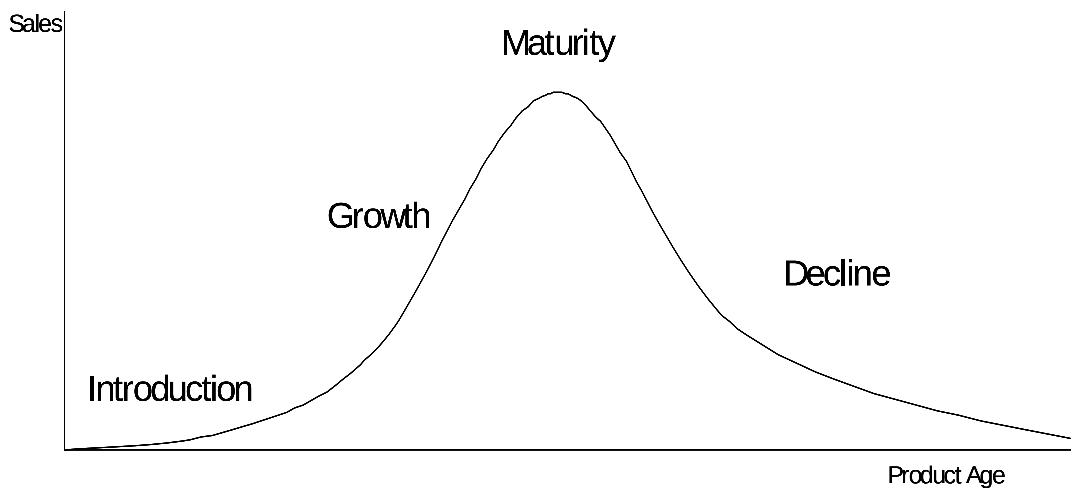

Figure 1: Product Life Cycle

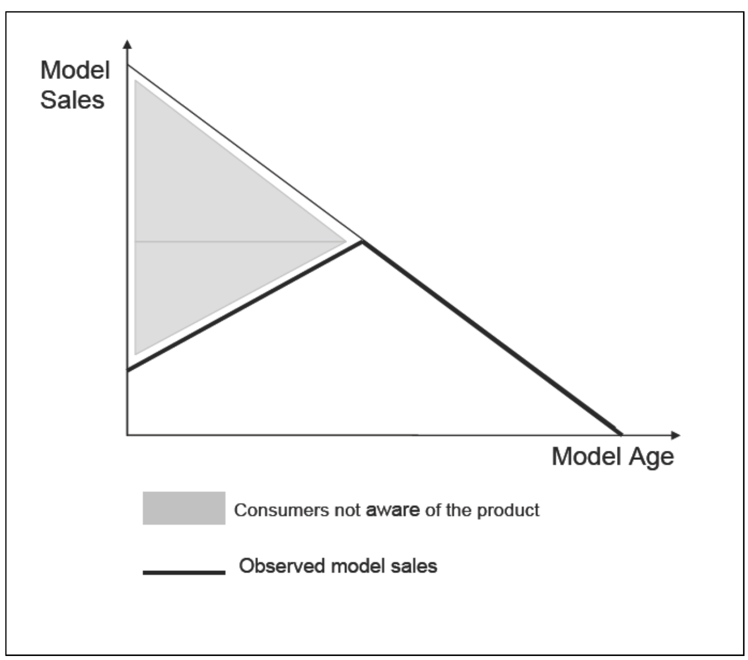

Figure 2: Information Diffusion of a New Product.

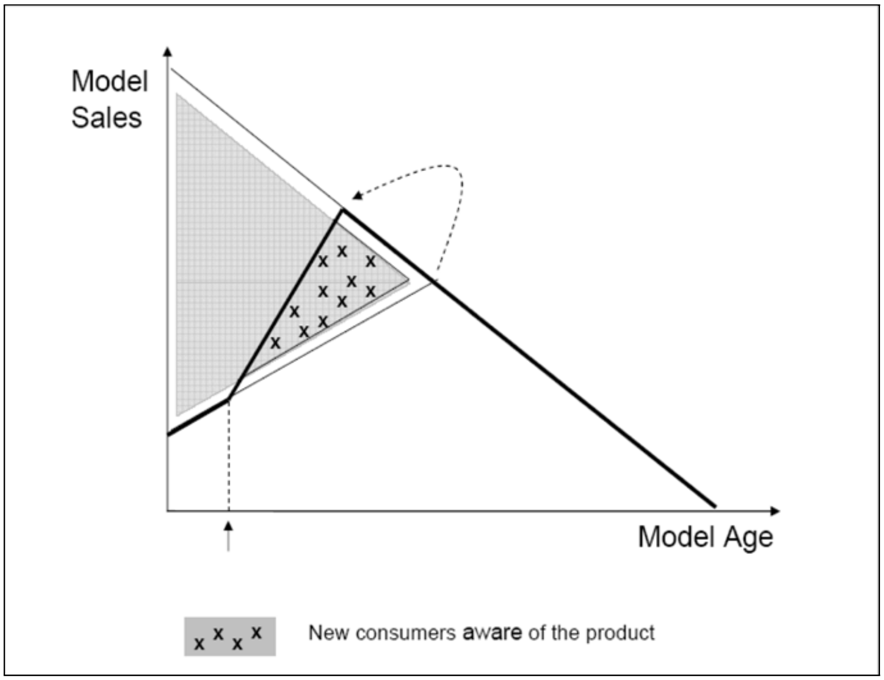

Figure 3: Impact of Advertising in the Future Consumers Aware of Product.

Figure 4: Estimated age effect on market share (ln(sj)-ln(s0)), using IV Logit demand

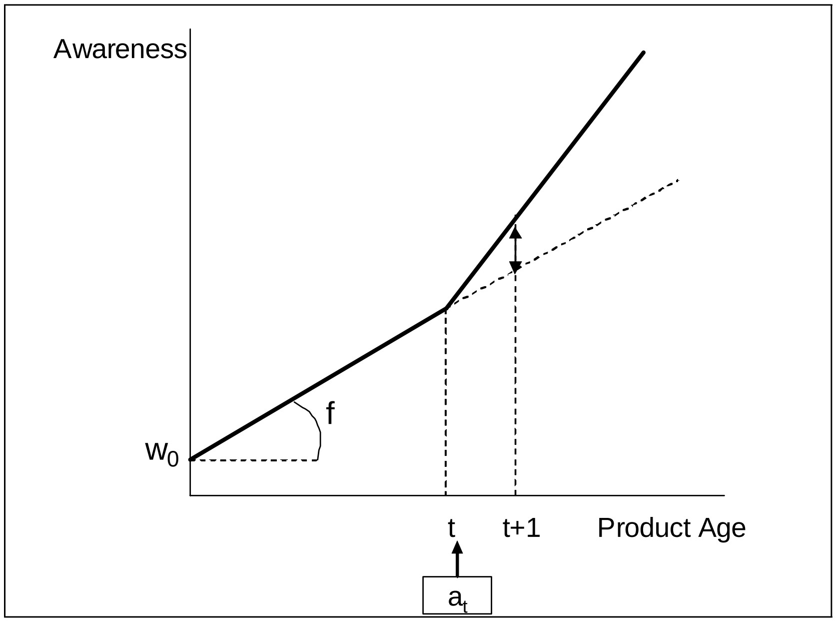

Figure 5: Product Awareness Level

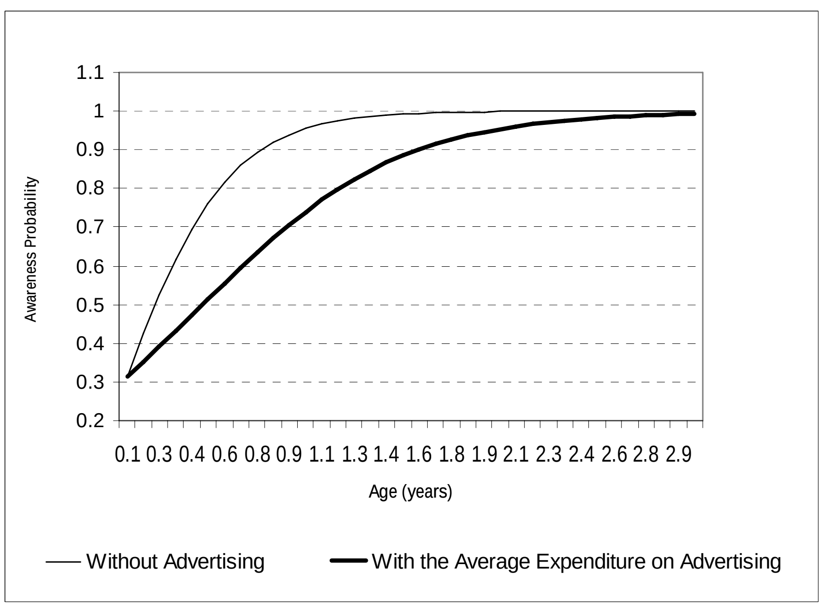

Figure 6: Awarenes Probability