Quantitative Marketing and Economics, 3, 107–144, 2005.

An Empirical Model of Advertising Dynamics

JEAN-PIERRE DUBÉ*
GÜNTER J. HITSCH
PUNEET MANCHANDA
University of Chicago, 5807 S. Woodlawan Avenue, Chicago, IL 60637
E-mail: jdube@gsb.uchicago.edu
E-mail: guenter.hitsch@gsb.uchicago.edu
E-mail: puneet.manchanda@ChicagoGSB.edu

Abstract. This paper develops a model of dynamic advertising competition, and applies it to the problem of optimal advertising scheduling through time. In many industries we observe advertising "pulsing", whereby firms systematically switch advertising on and off at a high-frequency. Hence, we observe periods of zero and non-zero advertising, as opposed to a steady level of positive advertising. Previous research has rationalized pulsing through two features of the sale response function: an S-shaped response to advertising, and long-run effects of current advertising on demand. Despite considerable evidence for advertising carry-over, existing evidence for non-convexities in the shape of the sales-response to advertising has been limited and, often, mixed. We show how both features can be included in a discrete choice based demand system and estimated using a simple partial maximum likelihood estimator. The demand estimates are then taken to the supply side, where we simulate the outcome of a dynamic game using the Markov perfect equilibrium (MPE) concept. Our objective is not to test for the specific game generating observed advertising levels. Rather, we wish to verify whether the use of pulsing (on and off) can be justified as an equilibrium advertising practice. We solve for the equilibrium using numerical dynamic programming methods. The flexibility provided by the numerical solution method allows us to improve on the existing literature, which typically considers only two competitors, and places strong restrictions on the demand models for which the supply side policies can be obtained. We estimate the demand model using data from the Frozen Entree product category. We find evidence for a threshold effect, which is qualitatively similar to the aforementioned S-shaped advertising response. We also show that the threshold is robust to functional form assumptions for the marginal impact of advertising on demand. Our estimates, which are obtained without imposing any supply side restrictions, imply that firms should indeed pulse in equilibrium. Predicted advertising in the MPE is higher, on average, than observed advertising. On average, the optimal advertising policies yield a moderate profit improvement over the profits under observed advertising.

Key words. advertising, dynamic oligopoly, Markov perfect equilibrium, pulsing

JEL Classification: L11, L66, M30, M37, R12

1. Introduction

We investigate how firms should optimally schedule their advertising over time. In addressing this issue, we consider the role of the response of sales to advertising. The topic of optimal dynamic advertising scheduling has a long history in the academic marketing literature, and is of course relevant to marketing practitioners. Observed advertising in many industries suggests that advertising scheduling is indeed a dynamic problem, and cannot be addressed in a purely static model. For example, many consumer packaged goods (CPG)

*Corresponding author.

companies use “pulsing” strategies, whereby advertising is concentrated in few weeks, followed by periods of zero advertising. This particular form of advertising is well-known in the marketing literature (Lilien et al., 1992, chapter 6). The theoretical rationale for pulsing is mostly based on an S-shaped sales response to advertising. The marginal return to advertising is increasing for small advertising levels below an inflection point. Above this region, marginal returns are decreasing. Hence, firms either choose a positive level of advertising in the region of decreasing returns, or do not advertise at all. Therefore, if advertising has long-run effects on demand, it can be optimal to bunch all advertising in a few periods.

While the advertising literature has addressed the practice of pulsing, there has been a marked disjoint between the theoretical and the empirical work in this area. The existing theoretical literature has shown that under certain conditions pulsing is an optimal advertising strategy. Mainly in the context of monopoly, the literature has identified the role of demand-side factors, in particular an S-shaped demand or sales response to advertising, as a source of pulsing (Sasieni, 1971; Simon, 1982; Mahajan and Muller, 1986, Feinberg, 1992).[1] Villas-Boas (1993) extends this literature to allow for competition. Also under the assumption of an S-shaped sales response, he shows that his model exhibits equilibria where firms use “alternating strategies,” such that they pulse out of phase. The empirical literature has mostly focused on detecting an S-shaped sales response to advertising (e.g. Simon, 1969; Rao and Miller, 1975; Wittink, 1977; Vakratsas et al., 2004). The evidence for an S-shape is mixed. Furthermore, many of the estimated sales-response models would not generate pulsing as an optimal advertising strategy. Hence, the question whether firms should actually pulse, given estimated product demand, has been largely unaddressed. The notable exception is Simon (1982), who estimates a monopolistic demand equation that is associated with an optimal pulsing policy. Simon’s demand system, however, is very restrictive, and he does not allow for competition.[2] Omitting competitive variables from a demand function is highly restrictive, and limits the usefulness of the resulting policy recommendations.

The solution of dynamic advertising problems under competition has been considered by a large literature that is based on differential game methods (Erickson, 1995, 2002), and a smaller literature that considers linear-quadratic games (Slade, 1995). Both approaches admit analytic solutions of the equilibrium advertising strategies. However, these solutions can only be obtained for demand systems that are very restrictive for empirical

1 Alternative theories have also been proposed. Bronnenberg (1998) derives pulsing in the context of a monopolist facing a Markovian sales-response function. Naik et al. (1998) consider copy wearout as a source of pulsing. In their model, the effectiveness of advertising decreases through time due to copy wearout. Once a firm stops advertising, the effectiveness of copy “recovers,” and the firm will start to advertise again after a certain period of time. Of course, this theory is based on the assumption that firms are committed to a specific copy. In other words, they cannot alternate between different copy over time to avoid wearout. Since our data does not include information about advertising copy or creative, we do not consider wearout as a source of pulsing. Furthermore, discussions with industry experts revealed that copy strategy is probably of only minor importance in the Frozen Entrée category.

2 Simon uses the demand specification $Q_t = \alpha + \lambda Q_{t-1} + \beta \log(A_t) + \gamma \max(0, A_t - A_{t-1})$, where Q_t denotes sales, and A_t is advertising. Note that this demand model precludes ‘true’ pulsing, i.e. an alternation between positive and zero advertising levels. Instead, Simon focusses on the alternation between ‘small’ and ‘large’ advertising levels, which he interprets as pulsing.

applications. In particular, an S-shaped sales response cannot be incorporated in these demand systems, which precludes a demand side explanation of pulsing as an optimal advertising strategy.[3,4]

This paper resolves the aforementioned shortcomings by using computational methods based on numerical dynamic programming to calculate optimal advertising strategies.[5]. It is not feasible to solve analytically for the equilibrium in the context of the model we describe below due to the complex demand system used. Our method places few restrictions on the demand side, and has the potential to handle most discrete choice based demand systems, such as a logit, random coefficients logit, or probit models. Hence, we can consider classes of demand models that are compatible with those used in the empirical literature on consumer choice.[6]

Our approach has two steps. First, we estimate product demand without imposing any supply-side restrictions such as the assumption that observed advertising was generated 'optimally', or from a particular model of firm conduct. Second, we solve for the optimal advertising policies implied by the estimated demand and profit functions.[7] Since it is difficult to provide a general proof for the incidence of pulsing we instead use this calibration to verify that pulsing would arise given our empirically estimated demand system. To capture the competitive nature of the oligopoly setting in our data, we specify a dynamic game and solve for the set of Markov Perfect Equilibrium advertising strategies. In this respect, the numerical solution techniques we use are essentially dynamic programming techniques applied to the case of strategic interaction (Pakes and McGuire, 1994).

On the demand side, we estimate a multinomial logit demand system. Our specification of the effect of brand advertising on demand allows us to carry-out several tests. We show that if the response of advertising were globally concave or linear, our model would not predict pulsing (unless we included a random state variable to shift the marginal return to advertising between periods). Hence, we test whether the sales-response to advertising exhibits a non-convexity in the form of a threshold below which advertising has zero impact. We diagnose the validity of the threshold effect using a non-parametric spline function. Our specification also allows us to test whether the impact of current advertising exhibits carry-over effects on future demand, which we model as a "goodwill" stock. Evidence for both a threshold effect and a carry-over effect would imply that the goodwill stock is a non-linear distributed lag of advertising, in contrast with the extant literature.

3 Other approaches, in particular "conjectural variations methods" have been used to study advertising dynamics; this literature, however, is mostly concerned with testing for particular modes of conduct, without actually being able to solve for the equilibrium advertising policy (Samuelson and Roberts, 1988; Vilcassim et al., 1999).

4 Typically, the differential games literature does not consider marketing instruments apart from advertising, such as pricing. Also, analytic solutions of the market equilibrium are only available for the case of duopoly markets.

5 A related empirical literature has used similar numerical dynamic programming tools to study consumer choices under forward-looking behavior (e.g., Erdem and Keane, 1996).

6 In a purely theoretical context, similar numerical tools have been used by Doraszelski and Markovich (2003) to solve for the long-run impact of advertising on industry structure, where the role of pulsing is not addressed.

7 A similar approach has been used to study the dynamics of pricing, entry and exit in the commercial aircraft industry (Benkard, 2001).

We use the demand side estimates to explore the implications for equilibrium advertising in the context of the oligopolistic industry structure observed in the data. The intertemporal effects of advertising on demand give rise to a dynamic game. The solution concept used is Markov perfect equilibrium (MPE). In this equilibrium, firms base their advertising decisions on the realizations of the payoff-relevant state variables, which, in our example, are the goodwill levels for each firm.

The structural parameters of our model are estimated using Scantrac level scanner data for the Frozen Entrée product category. The data comprise aggregate weekly sales, prices, and advertising levels for each brand in 18 major US city markets. Of particular importance are our advertising data. Since advertising is reported weekly, we can observe the "pulses" over time. In contrast, pulsing would not be observed in more typical time-aggregated advertising databases (e.g., quarterly or annual). Furthermore, advertising is measured in Gross Rating Points (GRPs), as opposed to more typical data sources reporting advertising expenditures. The fact that many companies sign advertising contracts with advertising publishers based on delivered GRP levels makes GRPs a more realistic unit of analysis. GRPs capture the typical exposure of households to advertising in each market/week, which provides a more accurate depiction of advertising effort than the dollar level of advertising expenditures. Another advantage of this data is the long 155-week time series. The long panel allows us to estimate market-specific product fixed-effects, to help avoid endogeneity bias. Moreover, the long time-series is key for measuring the long-run "carry-over" effects of advertising. A limitation of our aggregate data is that we cannot make definitive conclusions about the microeconomic implications of advertising, such as the role of the creative or the distinction between informative and persuasive effects of advertising. Hence, we remain agnostic about the precise structural interpretation of our estimated advertising response function.

On the demand-side, we find strong evidence of a threshold effect, whereby advertising has no significant impact below the threshold level. To check the robustness of this result, we experiment with a flexible spline approximation of the advertising response. The results indicate a threshold of about 32 GRPs. We also find strong evidence of advertising carry-over effects, i.e., goodwill levels depreciate slowly over time. Our findings for carry-over are comparable to those documented in the meta analysis of Assmus, Farley and Lehmann (1984). In this respect, our work contributes to a large literature measuring the long-run effects of advertising on sales using aggregate data (Clarke, 1976).

Next, we use the demand estimates to analyze optimal advertising strategies in a dynamic oligopoly. We solve for the optimal advertising strategies using our dynamic programming algorithm, based on the demand side estimates. Our discussions with firms indicate that actual advertising policy uses heuristics similar in spirit to the state-dependent MPE policies. Nevertheless, we do not view our MPE simulations as capturing the true advertising rules of firms. We use the MPE purely as a basis of comparison to assess the rationality of the observed advertising practices in the data. We find that pulsing is indeed the optimal advertising strategy in our empirical application. Therefore, optimal advertising exhibits patterns that resemble the observed advertising patterns in the Frozen Entrée category. We attribute the finding of optimal pulsing to the shape of demand. We estimate several alternative demand systems that relax some of our conditions on the sales-response to advertising. Each of these models yields inferior in-sample fit. None of these comparison models give rise to pulsing as an optimal strategy. Hence, we find that the proposed model provides superior in-sample fit and generates advertising patterns that are more consistent with observed practices. Interestingly, our results suggest that firms should optimally advertise more, on average, than observed. In particular, firms should optimally advertise at a higher frequency than in the data. We also find that predicted optimal advertising leads to a moderate profit improvement over the profitability under the observed advertising levels. Finally, almost all firms would have lower profits if they did not advertise at all.

The remainder of the paper is organized as follows. We conclude the introduction with a brief discussion of advertising scheduling in practice. In section two, we outline our general framework to model the effect of current and past advertising on demand, and the resulting supply-side model of optimal advertising and pricing. In section three, we discuss the specifics of the empirical demand system used in this paper along with the estimation approach. In section four we present the data, estimation results, and supply-side predictions. Conclusions and a general discussion of results and future research appear in section five.

1.1. Pulsing strategies in practice

The widespread practice of pulsing has also been a controversial topic in industry since the 1970s. Early pulsing advocates justified this practice as an efficient means of saving money with a limited media budget (Kingman, 1977). The practice was formalized as *effective frequency planning* (EFP) (Naples, 1979), which focused ad strategy primarily on frequency (the number of times an individual sees an ad) rather than on reach (the number of individuals in a target population who view the ad). The EFP approach was predicated on the notion that advertising response is subject to a threshold effect and diminishing returns, reflected in an S-shaped response curve. For instance, Krugam (1972) documented experimental evidence that three exposures were required to generate a response to advertising. Pulsing was recommended as an effective way for advertisers to generate sufficient frequency in ad exposures while managing a limited advertising budget.[8] Essentially, concentrating ads into flights generates relatively more exposures but, in contrast with a continuous level of advertising, less reach. By 1994, 68.3 percent of media-planners in the top 100 US advertising agencies were using EFP (Leckenby and Kim, 1994), despite the on-going empirical academic debate as to whether advertising response is indeed S-shaped versus concave (see Cannon et al., 2002 for a discussion) and whether a threshold effect exists (Jones, 1995). Interestingly, a related literature has also documented similar threshold effects for customer service. The implications of thresholds for customer service are similar in spirit to EFP as optimal service strategies involve serving fewer customers (Shugan and Radas, 2003).

More recently, pulsing has re-emerged as a recommended effective advertising practice in response to the measurement of adstock, or carry-over effects (see Ephron, 2002 for a discussion). In this line of reasoning, if media-planners believe advertising effect decays

8 In theory, one could also motivate pulsing if there were fixed costs associated with running any given ad campaign. In the model, this would amount to paying a fixed cost as well as a per-GRP cost in any week with positive GRP levels. Since the existing literature does not address this type of ad buying institution, we do not address it in the paper.

slowly over time, then pulsing can be an effective cost-saving device. Our research helps resolve this debate by studying theoretically the optimality of pulsing under fairly general conditions. We also test empirically whether the various characteristics of consumer response to advertising needed for pulsing are supported by our data.

2. A model of dynamic advertising competition

This section lays out the details of our dynamic advertising oligopoly model, and the associated Markov perfect equilibrium concept. We consider a market with J competing firms. Each firm produces one product, sometimes called a "brand". Firms compete, week after week, in advertising and prices. Time is discrete, $t = 0, 1, \ldots$ We first describe how advertising influences product demand. Then we discuss how firms choose their optimal product prices and advertising levels in order to maximize their objective function, the expected present discounted value of profits. Finally, we define the Markov perfect equilibrium solution concept that captures the strategic interactions between firms.

2.1. Advertising

The dynamics in our model stem from *advertising carry-over*, i.e. the long-term consequences of advertising on demand. Current advertising may affect product demand in the future. We formalize this intertemporal dependence through a *goodwill stock*, denoted by g_{jt}. In previous research, goodwill has been defined as a distributed lag of advertising, $g_{jt} = \sum_{k=1}^{L} \lambda^k A_{j,t-k}$. However, as we will explain in detail below, this formulation cannot account for pulsing because advertising enters goodwill linearly. Instead, we specify a model in which goodwill is a distributed lag of a non-linear transformation of advertising.

Firms use advertising to increase the beginning of period goodwill stock and create the *augmented goodwill stock*

$$g_{jt}^a = g_{jt} + \psi(A_{jt}). \tag{1}$$

ψ can be thought of as a "goodwill production function". In general, ψ is non-linear. We assume that $\psi(0) = 0$, and that ψ is non-decreasing in A. Some possibilities for the specification of the functional form of ψ will be discussed below. Augmented goodwill enters product demand. Over time, augmented goodwill in a period t depreciates stochastically and becomes the beginning of period goodwill stock in period $t + 1$

$$g_{j,t+1} = \lambda g_{jt}^a + \nu_{j,t+1} = \lambda(g_{jt} + \psi(A_{jt})) + \nu_{j,t+1}. \tag{2}$$

We assume that $0 < \lambda < 1$, such that goodwill depreciates in expectation, and that ν_{jt} is iid across time periods. We include the error term, ν, to capture the possibility that some aspects of advertising are not captured by the data. For instance, while GRP levels capture the intensity (frequency and reach) of advertising each week, they do not capture idiosyncrasies of the ads themselves. For example, variations in the effectiveness of the advertising copy, or variations in the composition of the audience reached may lead to different rates at which an ad is remembered and impacts on demand through time. Expanding equation (2), we find that

$$g_{jt} = \sum_{k=1}^{L} \lambda^k \psi(A_{j,t-k}) + \lambda^{L+1} g^a_{j,t-L-1} + \omega_{jt}, \tag{3}$$

where $\omega_{jt} \equiv \sum_{k=0}^{L} \lambda^k \nu_{j,t-k}$. This expansion shows that we can think of goodwill as a distributed lag of a (non-linear) transformation of advertising.

2.2. *Product demand and profits*

Product demand for firm j is a function of all firms' goodwill levels, advertising, prices, and a vector of demand shocks:

$$Q_{jt} = Q_j(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t, \boldsymbol{\xi}_t).$$

More specifically, demand is a function of augmented goodwill only, i.e. different levels of goodwill and advertising that produce the same amount of augmented goodwill also have the same effect on product demand. $\mathbf{g}_t = (g_{1t}, \ldots, g_{Jt})$, and the advertising, price, and demand shock vectors are defined in the same way. The per-period profit flow for firm j is

$$\tilde{\pi}_j = (P_{jt} - c_j) Q_j(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t, \boldsymbol{\xi}_t) - k A_{jt}.$$

c_j is the unit cost of production, and k is the cost of delivering one gross rating point. Prices and advertising are set before the demand shocks ε are realized, hence only expected profits matter for the firms' choices. Expected per-period profits are defined as

$$\pi_j = \pi_j(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t) = \int (P_{jt} - c_j) Q_j(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t, \boldsymbol{\xi}) p(\boldsymbol{\xi})\, d\boldsymbol{\xi} - k A_{jt}. \tag{4}$$

In this definition we implicitly assume that $\boldsymbol{\xi}_t$ is iid.

2.3. *Firms' decisions*

The exact timing of the advertising game is as follows. At the beginning of each period, the state of the market, denoted by $\mathbf{g}_t$, is observed by all firms. Based on the observed state, firms make their marketing decisions $\sigma_j(\mathbf{g}_t) = (A_{jt}, P_{jt})$. In principle, the state vector could include the entire past history of the game, i.e. past prices, advertising, sales, etc. However, we restrict our attention to games where firms' decisions are only based on payoff-relevant state variables. Given our formulation of product demand, the current goodwill levels provide all necessary information to forecast current and future sales. Hence, the state vector $\mathbf{g}_t$ contains all the payoff-relevant information. Once the state vector has been realized, and firms have made their pricing and advertising decision, the demand shocks $\boldsymbol{\xi}_t$ are realized, and the firms receive their current period profits.

A strategy profile $\sigma = (\sigma_1, \dots, \sigma_J)$ lists the decision rules of all firms. The expected present discounted value (PDV) of profits for firm j under the current state $\mathbf{g}_t$ and the strategy profile σ is

$$V_j(\mathbf{g}_t \mid \sigma) = \mathbb{E}\left[\sum_{s=t}^{\infty} \beta^{s-t} \pi_j(\mathbf{g}_s, \sigma_j(\mathbf{g}_s)) \mid \mathbf{g}_t\right]. \tag{5}$$

Firm j chooses a strategy σ_j, i.e. a sequence of state dependent advertising and price levels, in order to maximize $V_j(\mathbf{g}_t|\sigma)$. To calculate the expectation (5), the firm needs to know the evolution of state variables $\mathbf{g}_t$, and thus marketing decisions $\sigma_j(\mathbf{g}_t)$. From equation (2), we see that the goodwill stock g_{jt} follows a Markov process with transition density $p(\cdot \mid g_{jt}, A_{jt})$. Hence, because the depreciation shocks ν_{jt} are iid, the state vector has a Markov transition density

$$p(\mathbf{g}_{t+1} \mid \mathbf{g}_t, \mathbf{A}_t) = \prod_{j=1}^{J} p(g_{j,t+1} \mid g_{jt}, A_{jt}). \tag{6}$$

By assumption, firms make their pricing and advertising decisions only dependent on the information contained in the current state vector. This assumption also rules out time dependent strategies. Denote firm j's *Markov strategy* by $\sigma_j : \mathbf{g} \rightarrow \sigma_j(\mathbf{g}) = (A_j(\mathbf{g}), P_j(\mathbf{g}))$. Firm j makes an assumption on the competitive strategy profile $\sigma_{-j} = (\sigma_1, \dots, \sigma_{j-1}, \sigma_{j+1}, \dots, \sigma_J)$, and chooses its own decision rule σ_j. Together with the transition density of the state vectors, (6), the expectation (5) is then well defined.

2.4. *Equilibrium*

As in a dynamic programming problem without any strategic interaction, equilibrium decisions are described by a value function, one for each firm, that satisfies the *Bellman equation*,

$$V_j(\mathbf{g} \mid \sigma) = \sup_{\alpha \in \mathbb{R}_+^2} \left\{ \pi_j(\mathbf{g}, \alpha, \sigma_{-j}(\mathbf{g})) + \beta \int V_j(\mathbf{g}' \mid \sigma) p(\mathbf{g}' \mid \mathbf{g}, \alpha, \sigma_{-j}(\mathbf{g}))\, d\mathbf{g}' \right\}. \tag{7}$$

The supremum is taken with respect to $\alpha = (A_j, P_j)$, an advertising level and price set by firm j. The Bellman equation is defined conditional on a specific competitive strategy profile σ_{-j}, i.e. a specific guess about the behavior of the firm's competitors. The right-hand side of the Bellman equation defines the best response to σ_{-j}. A Markov perfect equilibrium (MPE)[9] of the dynamic game is a list of strategies, $\sigma^* = (\sigma_1^*, \dots, \sigma_J^*)$, such that no firm deviates from the action prescribed by σ_j^* in any subgame that starts at some state $\mathbf{g}$.

9 See Maskin and Tirole (2001) for a concise treatment of the MPE concept.

Definition. A Markov perfect equilibrium is a Markov strategy profile σ^* such that

$$V_j(\mathbf{g} \mid \sigma^*) \geq \pi_j(\mathbf{g}, \alpha, \sigma^*_{-j}(\mathbf{g})) + \beta \int V_j(\mathbf{g}' \mid \sigma^*) p(\mathbf{g}' \mid \mathbf{g}, \alpha, \sigma^*_{-j}(\mathbf{g})) d\mathbf{g}'$$

for all unilateral deviations $\alpha = (A_j, P_j)$, states $\mathbf{g}$, and firms j.

A MPE contains the basic idea of a Nash equilibrium—each firm maximizes its objective, the PDV of profits, given a belief about its competitors' behavior, and the beliefs are mutually consistent. The equilibrium is perfect, i.e. constitutes a Nash equilibrium for each subgame that starts at some realization of the state vector. The important aspect of the MPE concept is that strategies are, relatively speaking, 'simple'. In principle, each firm could make its actions dependent on the whole history of the game, in particular the past actions chosen by all firms. In a MPE, however, strategies are only dependent on payoff relevant variables. The current goodwill vector describes all the payoff relevant information. Furthermore, if firm j's competitors make their decisions only based on $\mathbf{g}_t$, firm j cannot improve its profits by making its current advertising and pricing decisions dependent on any extra information, which is evident from (7). Note that we restrict our attention to pure strategies. A more general model would allow for mixed strategies, i.e. probability distributions on all possible advertising levels and prices. The numerical solution of a model with mixed strategies, however, would be substantially more difficult.

A MPE in pure strategies of the specific dynamic game considered in this paper need not exist, and if it exists, it need not be unique. General conditions for the existence of equilibria in games that are similar to the one considered here are given in Ericson and Pakes (1995) and Doraszelski and Satterthwaite (2003). However, due to differences in the model setup, the results of these authors cannot be directly applied to our situation. What is relevant for us is whether an equilibrium exists at the estimated parameter values. Existence for a specific version of the model is checked automatically by our numerical solution algorithm. If this algorithm converges, the existence of an equilibrium is established. Note that we do not impose any supply side restrictions on our estimation procedure. Therefore, the existence or uniqueness of an equilibrium is irrelevant for the validity of the estimator.

2.5. *Discussion*

As stressed before, the MPE concept involves relatively simple strategies in a complex dynamic decision problem. We tried to assess whether these particular strategies, by which current advertising is based on an appropriately defined average of past own and competitive advertising, bear any resemblance to the way advertising is set in the Frozen Entrée category, as reported by the industry participants. To that aim, we spoke with several industry experts regarding the institutional details of setting a media plan.[10] Our discussions indicated that firms do adjust their marketing instruments on a periodic basis in this category (and in general in consumer packaged goods industries) in response to changes in the

10 We are very grateful to Dennis Bender, Mike Duffy and Suresh Ramanathan for their discussions of media-planning and competitive advertising strategy.

market. Specifically, managers track their own and their competitors' advertising efforts using data similar to those described below. In a given week, adjustments to a media plan may be made in response to recent own and competitive advertising levels. For instance, managers understand that recent heavy competitive advertising can dilute own advertising effort, requiring additional advertising as a counter-measure. The notion of Markov perfect strategies captures this institutional behavior, with the additional assumption that managers understand how to convert own and competitive advertising levels into goodwill stocks. Hence, we conclude that the firms in the Frozen Entrée category advertise in a manner that is qualitatively consistent with a Markov perfect equilibrium. The finding that managers base current marketing decisions, to some extent, on historic own and competitive marketing decisions is consistent with other academic surveys of managerial practice (Montgomery et al., 2005) as well as previous empirical research that attempts to estimate competitive reactions post-hoc (e.g., Leeflang and Wittink, 1992 and 1996).

3. The empirical specification and estimation strategy

Our objective is to estimate a demand system that incorporates the advertising response described in section two. At the estimation stage, we will test whether the demand system exhibits the threshold and carry-over conditions that give rise to pulsing. The demand system will then be used to calibrate the equilibrium model of advertising.

For the empirical analysis, we use the aggregate multinomial logit demand system. Generally, the supply-side calibration could accommodate virtually any other demand system, as long as the resulting profits can be expressed as in equation (4). We first specify the details of our demand model. Then we discuss some specifics of the demand model, and in particular under what demand conditions pulsing can arise as the optimal advertising strategy, and whether pulsing could arise even without an S-shaped advertising response. Next we introduce our partial maximum likelihood (PML) estimator, and conclude with a discussion of how our data can identify the model parameters.

3.1. Demand specification

A household (or consumer) h derives the following indirect utility from the consumption of product (brand) j:

$$U_{jt}^h = \alpha_j + \theta P_{jt} + \Gamma\left(g_{jt}^a\right) + \xi_{jt} + \varepsilon_{jt}^h.$$

α_j is a brand-specific intercept, that captures the utility of all observed and unobserved product attributes that do not vary over time. P_{jt} is the product price, and θ the marginal utility of income (i.e. the price-sensitivity parameter). Γ captures the utility effect of augmented goodwill. The reason why augmented goodwill does not enter utility in a linear form will become apparent below. Recall from above that the augmented goodwill, g_{jt}^a, is a function of the goodwill depreciation parameter, λ. Hence, we can construct a test for the carry-over effect of advertising by testing whether $\lambda = 0$. ξ_{jt} denotes temporal utility shocks that are observed by the consumers, but not the by the researcher. Finally, ε_{jt}^h is an idiosyncratic iid taste shock, that has a Type I extreme value distribution. In a given week, each household chooses to consume a serving of one of the brands, or does not participate in the market, i.e. chooses the outside alternative with utility $U_{0t}^h = \varepsilon_{0t}^h$.

Using the expression for logit choice probabilities, market demand is given by:

$$Q_{jt} = MS_{jt} = M \cdot \frac{\exp(U_{jt})}{1 + \sum_{k=1}^{J} \exp(U_{kt})} \tag{8}$$

where U_{jt} is the common component (across households) of indirect utility. M denotes the total market size, and the market share of each brand is S_{jt}.

The parametric specification for the advertising effect Γ is

$$\Gamma(g^a) = \begin{cases} \gamma \log(1 + g^a) & \text{if } g^a \geq 0, \\ 0 & \text{otherwise.} \end{cases} . \tag{9}$$

The rationale for this functional form, as opposed to a simple linear specification by which augmented goodwill enters mean utility, is as follows. The logit model, as other discrete choice models, such as a probit, exhibits an S-shaped relationship between the market share and the indirect utility from which the market share is derived. Typically, in markets with many products and a large outside share, each individual product commands only a small fraction of the market. Hence, its market share exhibits increasing returns to indirect utility. If goodwill entered indirect utility linearly, firms would have a very strong incentive to advertise and thus increase indirect utility. Instead, we choose the logarithmic form for Γ to provide a well behaved objective function and advertising optimization problem for each firm. It can be easily verified that the market share under this specification (8) is strictly concave in augmented goodwill, provided that $0 < \gamma \leq 1$.[11] On the other hand, the specification $\Gamma(g^a) = (g^a)^\gamma$ can lead to increasing returns to goodwill, even if $\gamma < 1$. At the estimation stage, we can test the relative fit of specifying $\Gamma(g^a)$ as linear versus as in the formulation in equation (9).

Next, we specify the form of the goodwill production function ψ. Whether firms optimally pulse or not depends crucially on the functional form of ψ. The basic intuition behind demand-based pulsing is the S-shaped response of demand to advertising. An extreme form of an S-shaped response is given by a threshold model, in which advertising is ineffective up to a specific level G. Beyond that threshold level, advertising has diminishing returns. In this setup, firms will either choose not to advertise at all, or to advertise at some level $A > G$. Because advertising has long-run effects on demand through goodwill, the marginal return to goodwill and hence advertising can be different across time. For example, at a small level of accumulated goodwill, the marginal return to advertising will be relatively large, and a firm may optimally choose a positive advertising level. Some of the created goodwill will still impact demand in the next period. Hence, the future marginal return to goodwill will be lower than the current marginal return. Therefore, a firm may no-longer find it optimal to advertise in the next period.

11 Note that $\partial S_j / \partial g_j^a = \Gamma'(g_j^a) S_j (1 - S_j) = \gamma(1 + g_j^a)^{\gamma - 1} \exp(\tilde{U}_j)/(1 + \sum_{k=1}^{J} \exp(U_k))$, where $\tilde{U}_j = U_j - \Gamma(g_j^a)$. Hence, the derivative is strictly decreasing in augmented goodwill if $0 < \gamma \leq 1$.

A simple, parametric form of ψ that includes a threshold effect is given by

$$\psi(A) = \begin{cases} \log(1 + A - G) & \text{if } A \geq G, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

G is a threshold parameter to be estimated. An attractive feature of this specification is that (10) nests the case of a purely concave sales response, $G = 0$. Hence, we can construct a parametric test of whether $\psi(A)$ is strictly concave or not by testing whether $G = 0$. Since the functional form of ψ is critical to our argument for pulsing, we want to exclude the possibility that our results are driven purely by functional form assumptions. Hence, we compare the parametric estimation results to a semi-parametric approach, where ψ is approximated by a B-spline:[12]

$$\psi(A) = \sum_{i=0}^{n} a_i N_i(A). \tag{11}$$

B-splines are convenient representations of spline functions, expressed as a linear combination of basis functions N_i. Each N_i is defined with respect to a knot sequence $(A_0, A_1, \ldots, A_{n+k})$. The control points a_j are parameters to be estimated. In the empirical application, we use basis functions of order three, which implies that the spline function consists of a succession of cubic polynomials on each knot interval. Roughly speaking, we use the B-spline estimates to construct a non-parametric test for the threshold effect. That is, we use them as a diagnostic tool to assess the validity of our specification (10).

3.2. Discussion

The proposed demand system, in particular the threshold in the goodwill production function, is sufficiently flexible to account for pulsing behavior in equilibrium based on demand-side factors. Whether pulsing will arise depends on the specific parameter values of the model. In particular, if the goodwill production ψ is globally concave, which is a special case of our model, pulsing cannot arise in a monopoly (see appendix A for a proof). We are unable to extend the proof of this statement to the case of oligopoly. However, we solved the model under oligopoly for various market structures (i.e., number of competing firms and demand specifications) and never obtained pulsing as an equilibrium outcome if the threshold in ψ was set to 0. We also note that a positive threshold per se is not sufficient to give rise to pulsing. For example, if the rate of goodwill depreciation is very fast (λ is very small), the marginal return to augmented goodwill, and hence advertising, may always be large enough to induce firms to advertise (i.e. firms advertise every week). On the other hand, if λ is sufficiently large, firms may not advertise in some periods because they still retain sufficient amounts of goodwill from the past, such that the return to advertising (augmented goodwill) does not exceed the cost.

12 See for example Chen and Conley (2001) for an application.

In summary, pulsing may but need not arise in our model. Whether firms should pulse is thus a matter of the particular parameter estimates obtained for a given application or market.

3.3. Partial maximum likelihood estimation

Based on the model described above, we now discuss the strategy used to estimate the demand parameters in the demand function (8). We do not impose any supply-side restrictions when we estimate demand. In this respect, our demand estimates are not predicated on assumed firm conduct. However, ignoring the supply-side could generate endogeneity biases that we discuss below.

Using a simple logarithmic transformation of (8), we obtain the following linear equation:

$$\begin{aligned} y_{jt} &\equiv \log(S_{jt}) - \log(S_{0t}) \\ &= \alpha_j + \theta P_{jt} + \Gamma(g_{jt}^a) + \xi_{jt}. \end{aligned} \tag{12}$$

From equation (3) we know how current goodwill is related to the initial augmented goodwill level, and the history of advertising and goodwill depreciation shocks:

$$g_{jt} = \sum_{k=1}^{t-1} \lambda^k \psi(A_{j,t-k}) + \lambda^t g_{j0}^a + \omega_{jt}, \qquad \omega_{jt} \equiv \sum_{k=0}^{t-1} \lambda^k \nu_{j,t-k}.$$

This expansion can be substituted for g_{jt}^a in the estimation equation. As discussed above, due to the S-shape of the logit market share equation, we do not expect to find a well-behaved advertising equilibrium if Γ is linear. However, if Γ is non-linear, then

$$m_{jt} = \alpha_j + \theta P_{jt} + \Gamma\left(\sum_{k=1}^{t-1} \lambda^k \psi(A_{j,t-k})\right)$$

is not the conditional expectation of y_{jt}, and hence a non-linear least squares estimator will not be consistent for the model parameters. Instead, we pursue a maximum likelihood based estimation approach.

Let $\mathbf{A}_j^t = (A_{j1}, \ldots, A_{jt})$ denote the sample history of observed advertising levels until time period t. y_{jt} is distributed according to

$$\begin{aligned} F_t\left(y \mid P_{jt}, \mathbf{A}_j^t; g_{j0}^a\right) &= \Pr\left\{y_{jt} \le y \mid P_{jt}, \mathbf{A}_j^t; g_{j0}^a\right\} \\ &= \Pr\left\{\alpha_j + \theta P_{jt} + \Gamma\left(g_{jt}^a\right) + \xi_{jt} \le y \mid P_{jt}, \mathbf{A}_j^t; g_{j0}^a\right\} \\ &= \Pr\left\{\xi_{jt} \le r(\omega_{jt}) \mid P_{jt}, \mathbf{A}_j^t; g_{j0}^a\right\}, \end{aligned} \tag{13}$$

where the function r in the last line is defined by

$$r(\omega_{jt}) = y - \alpha_j - \theta P_{jt} - \log\left(1 + \sum_{k=0}^{t-1} \lambda^k \psi(A_{j,t-k}) + \lambda^t g_{j0}^a + \omega_{jt}\right).$$

The conditional distribution of y_{jt} as expressed by the probability (13) can be calculated by integrating with respect to the joint distributions of ω_{jt} and ξ_{jt}. We assume that ξ_{jt} is normally distributed, with density $f_\xi = N(0, \sigma_\xi^2)$. ω_{jt} is a weighted sum of iid random variables, which are assumed to be normally distributed, $\nu_{jt} \sim N(0, \sigma_\nu^2)$. Hence, ω_{jt} is also normally distributed, with mean 0 and variance

$$\operatorname{var}(\omega_{jt}) = \sum_{k=0}^{t-1} \lambda^{2k} \sigma_\nu^2 = \frac{1 - \lambda^{2t}}{1 - \lambda^2} \sigma_\nu^2.$$

We denote its density by $f_{\omega t}$. Then

$$\Pr\left\{\xi_{jt} \leq r(\omega_{jt}) \mid P_{jt}, \mathbf{A}_j^t; g_{j0}^a\right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{r(\omega)} f_\xi(\xi) f_{\omega t}(\omega)\, d\xi\, d\omega.$$

Thus, the probability density of y_{jt} is given by

$$\begin{aligned} f_t\big(y \mid P_{jt}, \mathbf{A}_j^t; g_{j0}^a\big) &= \frac{d}{dy} F_t\big(y \mid P_{jt}, \mathbf{A}_j^t; g_{j0}^a\big) \\ &= \frac{d}{dy} \int_{-\infty}^{\infty} \int_{-\infty}^{r(\omega)} f_\xi(\xi) f_{\omega t}(\omega)\, d\xi\, d\omega \\ &= \int_{-\infty}^{\infty} f_\xi \Big(y - \alpha_j - \theta P_{jt} \\ &\quad - \log\Big(1 + \sum_{k=0}^{t-1} \lambda^k \psi(A_{j,t-k}) + \lambda^t g_{j0}^a + \omega_{jt}\Big) f_{\omega t}(\omega)\, d\omega. \end{aligned} \tag{14}$$

The final step in the calculation of the likelihood function handles the initial condition, g_{j0}^a. We assume that g_{j0}^a is normally distributed, with mean and variance

$$\begin{aligned} \mu_{gj} &= \frac{T^{-1} \sum_{k=1}^T \psi(A_{jt})}{1 - \lambda}, \\ \sigma_g^2 &= \operatorname{var}\left(\sum_{t=0}^{\infty} \lambda^t \nu_t\right) = \sum_{t=0}^{\infty} \lambda^{2t} \sigma_\nu^2 = \frac{1}{1 - \lambda^2} \sigma_\nu^2. \end{aligned}$$

The conditional density (14) involves the sum of two independent normal random variables, $\zeta_{jt} \equiv \lambda^t g_{j0}^a + \omega_{jt}$. Therefore, ζ_{jt} is itself normal with mean and variance

$$\begin{aligned} \mu_{jt} &= \lambda^t \mu_{gj} = \lambda^t \cdot \frac{T^{-1} \sum_{k=1}^T \psi(A_{jt})}{1 - \lambda}, \\ \sigma_{jt}^2 &= \lambda^{2t} \frac{1}{1 - \lambda^2} \sigma_\nu^2 + \frac{1 - \lambda^{2t}}{1 - \lambda^2} \sigma_\nu^2 = \frac{1}{1 - \lambda^2} \sigma_\nu^2. \end{aligned}$$

We can now derive the density of y_{jt}, conditional on the data $(P_{jt}, \mathbf{A}_j^t)$, and a vector Θ that collects all the model parameters:

$$\begin{aligned} & f_t(y \mid P_{jt}, \mathbf{A}_j^t; \Theta) \\ &= \int_{-\infty}^{\infty} f_t\big(y \mid P_{jt}, \mathbf{A}_j^t; g^a\big) f_{gj}(g^a) dg^a \\ &= \int_{-\infty}^{\infty} f_\xi \left(y - \alpha_j - \theta P_{jt} - \log \left(1 + \sum_{k=0}^{t-1} \lambda^k \psi(A_{j,t-k}) + \zeta \right) \right) f_{\zeta j}(\zeta)\, d\zeta. \end{aligned}$$

Note that the product of the individual densities, $\prod_{t=1}^{T} f_{jt}$, is not the joint density of the whole history of observations, $\mathbf{y}_j = (y_{j1}, \ldots, y_{jT})$. Due to the common initial condition, and the accumulated history of goodwill depreciation shocks, within-market observations of y_{jt} for a brand j are correlated. Hence, the current realization of y_{jt} provides information about future realizations of this variable. The joint density could be calculated by integrating over the whole history of goodwill depreciation shocks ν_{jt}. This is a computationally complex task, and would require us to calculate high-order integrals. Instead, we estimate Θ by a partial maximum likelihood (PML) estimator, that is obtained by maximizing the log-likelihood

$$l(\Theta) = \sum_{j=1}^{J} \sum_{t=1}^{T} \log f_t\big(y \mid P_{jt}, \mathbf{A}_j^t; \Theta\big). \tag{15}$$

The full MLE, on the other hand, would be based on the log-likelihood of the joint densities of $\mathbf{y}_j$, which would require us to calculate the aforementioned integrals of order T. Of course, this is not an impossible task, and it would be interesting to compare the two estimators in future research.

As discussed in Wooldridge (2001), the PML estimator provides a consistent estimate of the model parameter vector, Θ. This procedure is a special case of the more general quasi-maximum likelihood estimation approach that, under certain conditions, permits consistent parameter estimation with a misspecified likelihood (White, 1982). Since our data exhibit serial correlation, we choose a robust estimator for the asymptotic covariance matrix of the PML estimator. The asymptotic covariance matrix we use has the form $\mathbf{A}_0^{-1}\mathbf{B}_0\mathbf{A}_0^{-1}$, where $\mathbf{A}_0 = -\sum_{t=1}^{T} \mathbb{E}[\nabla_\Theta^2 l_i(\Theta_0)]$ and $\mathbf{B}_0 = \mathbb{E}[\nabla_\Theta l_i(\Theta_0) l_i(\Theta_0)']$ (Wooldridge, 2001, pp. 405–408).

3.4. Identification

A potential issue of concern is the identification of the advertising threshold. If firms are aware of the threshold, then we would only expect to observe advertising levels of zero or strictly greater than the threshold level. In this situation, it is unclear how the threshold level could be estimated by the estimation procedure. However, in our and many other situations this problem is avoided due to "make-goods," a common institutional practice in the advertising industry. Advertisers typically write contracts that specify the amount of gross rating points to be delivered in a given week. Delivered GRPs, however, depend on the actual realized reach and frequency of the aired ads. Often, the contracted amount of GRPs are not delivered entirely and, consequently, the television stations "make good" in subsequent periods by delivering the residual promised GRP levels. In our data, we frequently observe low GRP levels at the end of a flight of advertising. For instance, in Figure 1 we plot the advertising histories for Healthy Choice, Lean Cuisine and Stouffer's in Atlanta. For Lean Cuisine, between the fifth and seventh week, delivered GRPs are over 30. But in the eighth week, only 2.8 GRPs are delivered. Both Figure 1 and a histogram of all strictly positive advertising levels (Figure 2) reveal that there is substantial variation in ad levels in our data, and in particular that small advertising levels occur frequently. These aspects of the data are key to identifying the threshold and the overall shape of the goodwill production function ψ.

The data we use (described in the next section) contain time series observations for several brands in a cross section of geographic markets. Hence, it is important to control for differences in tastes and demographics, i.e. heterogeneity, across markets. We include a full set of market-specific brand intercepts in the demand equation. Since each market has 156 weeks per brand, we can estimate these intercepts as fixed-effects. The remaining parameters in the model capture time-variation in demand for a brand within and across markets. In particular, the estimated goodwill process will only account for deviations of realized market shares from market/brand specific mean market shares, and the relationship of these deviations to present and past advertising. Without fixed effects, the goodwill process might simply pick up heterogeneity across markets.

Another issue is whether an econometric endogeneity problem biases the estimated price elasticities and advertising effects. First, if firms observe the realizations of the unobserved attributes, ξ_{jt}, before they make their pricing and advertising decisions, the resulting dependence between the ξ_{jt}'s and the price and advertising levels will introduce endogeneity bias in our PML estimates. However, this problem is likely to be less severe in our application, because we estimate brand specific fixed effects separately in each market. Nevo (2001), for example, who faces a price endogeneity problem, controls for product fixed effects, but does not allow for these fixed effects to vary across markets due to the relatively short time series dimension of his data. Therefore, the concern is that systematic, time-invariant differences in the demand for a given brand across markets are correlated with systematic, time-invariant price differences. In contrast, in our application an endogeneity bias can only arise if weekly deviations of the ξ_{jt}'s from their market-specific means are correlated with prices or advertising levels. Firms, therefore, would need to be able to observe these shocks before they set their marketing variables. It is possible that this form of correlation is present in the data, for example due to couponing and promotion variables that are omitted from our demand system. But we believe that this problem is much less severe than the endogeneity problem due to persistent correlations between demand unobservables and prices that has been dealt with in the previous literature.

A second, related endogeneity problem concerns the relationship between advertising and the shocks to goodwill depreciation, ν_{jt}. In our model we assume that firms observe these shocks before they make their advertising decisions. If our data are in fact generated from our exact model, we face an advertising endogeneity problem. Negative shocks to goodwill will then be correlated with high advertising levels, and the estimated relationship

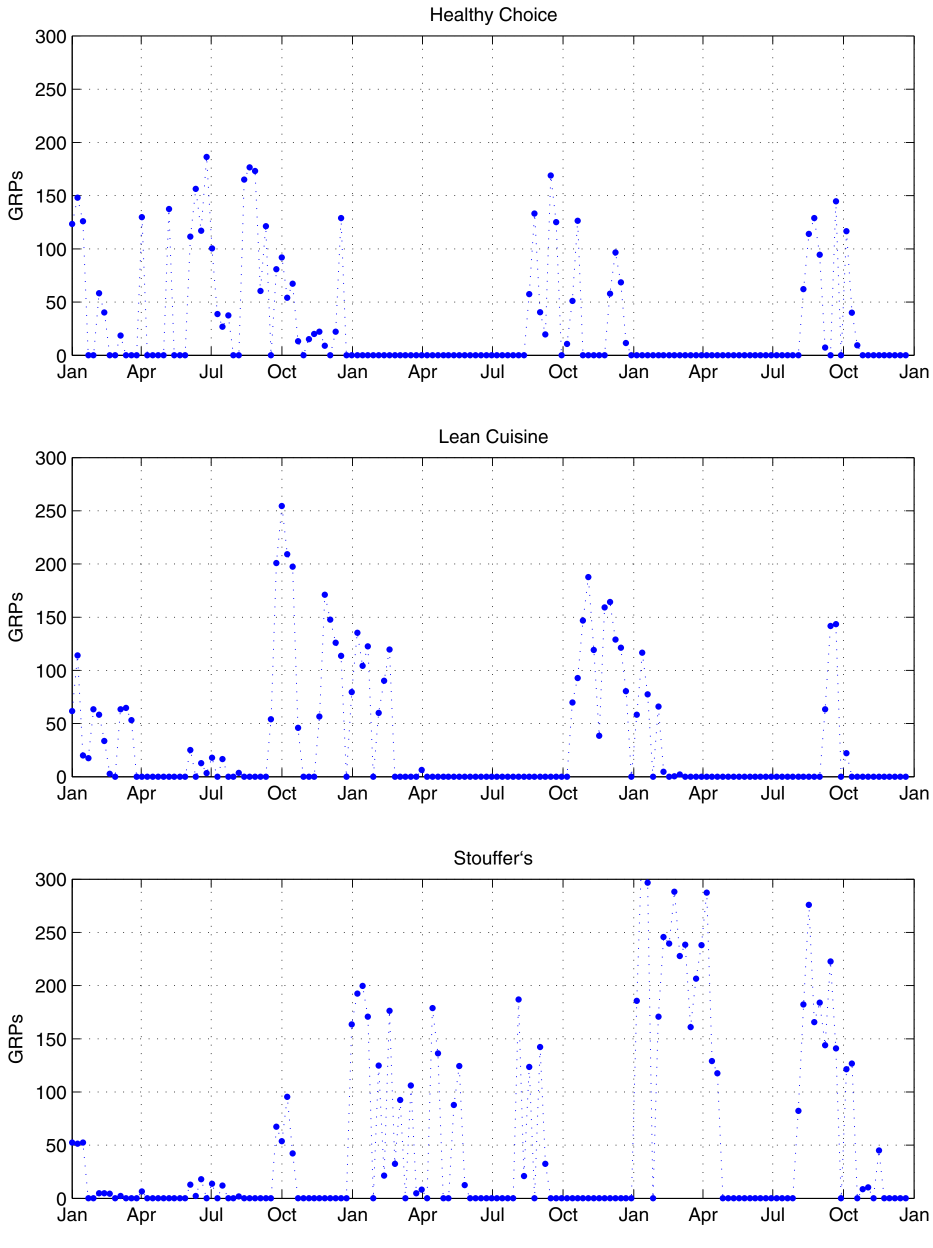

Figure 1. Advertising GRPs in Atlanta, 1991–1993.

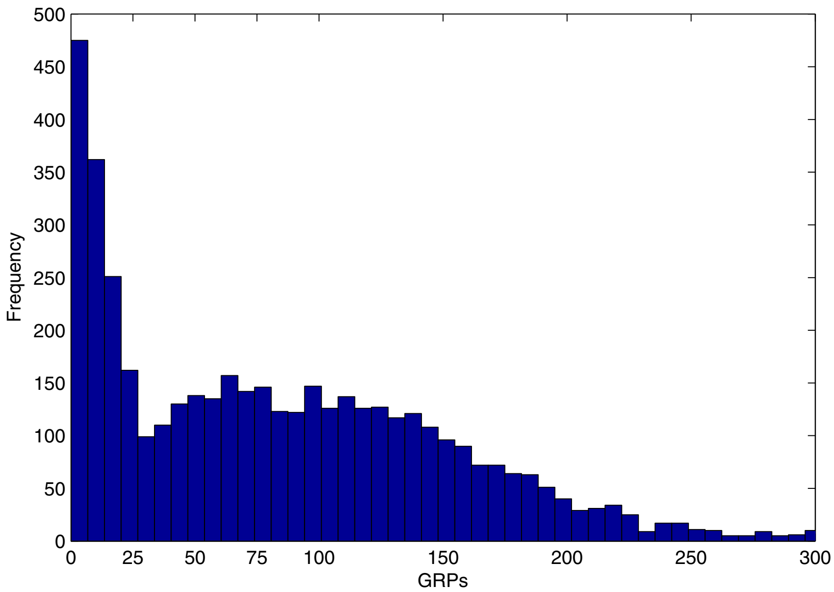

Figure 2. Frequency distribution of strictly positive GRP levels.

between demand and goodwill will be biased downwards. However, the institutional feature of make-goods is again helpful here, as it provides *exogenous* variation in advertising. Because make-goods are typically small advertising levels, we can consistently estimate the threshold in the advertising production function. Our main question, whether firms should pulse due to demand side factors, will therefore not be affected by an endogeneity problem. On the other hand, we might underpredict the optimal level of advertising, due to bias in the overall response of demand to goodwill, Γ.

For now, we acknowledge these two endogeneity issues. Endogeneity due to unobserved attributes has typically been approached in the literature using instrumental variable techniques (Berry et al., 1995). In the context of our ML based estimation framework, the analogous solution would require us to specify the joint distribution of the unobserved attributes, prices, and advertising levels (Villas-Boas and Winer, 1999), conditional on some instruments. With regards to advertising, this is a daunting task in our context due to the complicated mixed discrete/continuous distribution of advertising. Also, we have no good instruments for advertising available.[13] The previous literature has focused on price endogeneity, while the advertising endogeneity problem has mostly been neglected. An exception is Hitsch (2004), who solves the problem by imposing the optimal advertising policy on his ML estimator. This approach is of course not feasible in our application, as we want to know how firms should optimally advertise, and do not want to assume that observed advertising is already fully optimal.

13 We observe GRP costs only at the quarterly level, while our main data are at the weekly frequency.

A final issue is the identification of advertising carry-over effects. If firms advertised at an approximately constant level, sustaining a steady advertising rate across weeks, it would be difficult to measure carry-over. However, referring again to Figure 1, we see that in fact, due to pulsing, advertising is very volatile. The persistence of advertising is identified from the speed at which demand settles down to the zero-advertising level at the end of an advertising flight.

4. Results

In this section we discuss the data used for demand estimation, and our results. After describing the data from the Frozen Entrée category, we present our demand estimates. We then test the structure we impose on the formulation for the advertising effect on demand. Finally, our estimates are used to calibrate the dynamic advertising oligopoly model. We present the model's predictions for advertising and profits under two different assumptions on firm conduct, and compare the predictions to the observed advertising data and profits.

4.1. Data

The data used were collected by A.C. Nielsen and Competitive Media Research, and made available to us by Management Science Associates. The data consists of three years (January 1991 to January 1994) of weekly (156 weeks) sales and marketing mix variables across eighteen markets in the Frozen Entrée product category. The market definitions are based on A.C. Nielsen's SCANTRAK markets.

The Frozen Entrée market was worth $ 1.4 billion in retail sales (average across the period 1991–1994). The average volume growth was about 4–5% during this period.[14] There were no significant brand introductions or withdrawals during this period. The average expenditure on advertising was $ 127 million—thus the advertising/sales ratio was about 10%.[15] Five main brands—Budget Gourmet,[16] Stouffer's,[17] Swanson,[18] Healthy Choice,[19] and Lean Cuisine[20]—accounted for over 80% of the total sales in the category. Market shares are computed by assuming that each week, each individual in the market population chooses to consume a single serving of one of the brands, or does not participate in the category at all. The share of a brand is obtained by dividing total sales of a brand by the total population, i.e., total potential sales. The price variable is the average

14 See *Frozen Food Excecutive*, December, 1992–1994.

15 Data from *Leading National Advertisers*.

16 During this period, BG was marketed by the All American Gourmet Company, a subsidiary of Kraft. Kraft sold the BG brand to the Heinz in late 1994 (note that this is not reflected in our data since the last week in our data end on January 2, 1994). Heinz subsequently sold the brand to Luigino's Incorporated in 2001.

17 ST has been owned by Nestle since 1973.

18 Campbell Soup acquired SW in 1955. It was spun off as part of Vlasic Foods in 1998 and sold to Pinnacle Foods Corporation in 2001.

19 HC has been marketed by ConAgra since 1988.

20 Marketed by the Campbell Soup Co. during sample period. Campbell divested the brand after the end of the sample period.

Table 1. Descriptive statistics.

	Price ($)		GRPs		Share (%)	
Brand	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Boudget Gourmet	1.75	0.22	16.49	43.69	1.98	1.14
Healthy Choice	2.59	0.27	29.20	49.81	1.34	0.67
Lean Cuisine	2.20	0.25	34.93	54.39	1.46	1.15
Stouffer's	2.33	0.31	43.28	77.65	1.99	1.27
Swanson	2.00	0.20	9.76	40.78	1.68	0.83

weekly shelf price of a single-serving unit.[21] Descriptive statics of all variables appear in Table 1.

The advertising data consists of gross rating points (GRPs), representing the combined weight of advertising from different TV sources—network, spot, syndicated and cable.[22] GRPs are defined by the reach and frequency of the advertising in a given time period, where 'reach' is the fraction of the population that has been exposed to an ad at least once, and 'frequency' is the average number of times an exposed household has seen the ad. Hence, a GRP should be interpreted as the average number of times a typical household has been exposed to an ad in a given time period. With the exception of a few recent studies (e.g., Vilcassim et al., 1999), most prior research has used dollar advertising expenditures to capture the effects of advertising on sales. Advertising expenditures, however, are a measure of the cost of the advertising campaign, whereas GRPs more accurately represent the impact of the advertising on consumers' choice behavior. Furthermore, many companies sign advertising contracts with advertising publishers based on delivered GRP levels. In addition, GRPs measure advertising adjusted for the market size.

We now explore the observed advertising pulsing and flighting behavior in the data. In Figure 2, we provide a histogram of the observed positive GRP levels, pooled across all markets and brands. Roughly 60% of the weekly GRP levels are below 100, and 32% are between 100 and 200 rating points. To understand this wide range of GRP levels, we refer again to Figure 1, that shows the advertising history of the Lean Cuisine brand in Atlanta. We observe both weeks with large and small GRP levels, where the latter are often "make-good" weeks, during which the ad publisher provides extra GRPs when the previous week's target was not entirely fulfilled. Generally, the pulsing behavior is quite clear in the plots. That is, we observe distinct blocks of weeks without advertising followed by periods of on and off advertising.

An interesting feature of the data is the long panel structure which contains long histories of marketing information for several geographic markets. In Table 2, we decompose the variance of prices and shares across markets, brands and weeks. Looking only at main

21 Depending on the type of dish, the size of a single-serving varies between 8–11 oz. depending on the brand and main ingredient. Note that, for a given brand, the price is usually constant across single-servings.

22 Data from *Leading National Advertisers* shows that TV advertising dollars accounted for about 72% of all advertising dollars in this category during this period. Spot TV dollars represent about 50% of all TV advertising dollars and 35% of all advertising dollars.

Table 2. Analysis of variance of prices and shares.

	Market	Brand	Week	brand*market	brand*week
Price	0.13	0.56	0.02	0.72	0.67
Share	0.16	0.06	0.06	0.38	0.22

effects, brands seem to account for a large portion of price variation. However, we do not see very strong effects otherwise. Interestingly, we observe very substantial interaction effects between brands and markets and, to a slightly lesser extent, brands and weeks. The strong brand/market interaction indicates that geographic markets are quite different in nature. Also, despite the seemingly weak main effect of weekly variation, we see that there is considerable information in the time-series about specific brands.

In Table 3, we look at advertising patterns across markets. We compare the within-market mean GRP levels, conditional on advertising in a week, to assess the typical "pulse". We see some heterogeneity in the typical pulse for a brand overall. However, we see considerable dispersion in the pulse sizes for brands across markets. We also compare the within-market mean advertising frequencies to assess typical pulsing rates. We see that the typical pulsing rate is quite different across brands, as high as 43% of the weeks (in an average market) for Lean Cuisine, and as low as 13% for Swanson. Across markets, we see very little dispersion in the pulsing rates across markets for some brands, such as Lean Cuisine and Healthy Choice. For other brands, such as Stouffer's and Budget Gourmet, we see considerable dispersion in pulsing rates across markets. These results seem to indicate that firms do not run a uniform advertising plan across markets as we observe distinct differences both in the size and the frequency of pulses.

Table 3. Advertising patterns across markets.

	Mean	SD	Min	Max
Average positive GRP's				
Budget Gourmet	63.46	28.00	23.78	123.74
Healthy Choice	80.04	5.11	70.35	90.53
Lean Cuisine	81.84	13.14	62.05	103.58
Stouffer's	95.45	43.15	11.91	156.21
Average advertising frequency				
Swanson	71.10	26.47	21.76	109.53
Budget Gourmet	0.24	0.06	0.15	0.38
Healthy Choice	0.37	0.02	0.33	0.40
Lean Cuisine	0.43	0.02	0.37	0.47
Stouffer's	0.39	0.15	0.11	0.55
Swanson	0.13	0.04	0.08	0.21

Note: The Table displays summary statistics of within market average positive advertising levels and advertising frequencies across all 18 markets.

Table 4. Demand estimates for logit demand system.

	Threshold Model		Spline Model	
	Coefficient	s.e.	Coefficient	s.e.
Log-likelihood	171.2434		179.9295	
BIC	574.28		690.60	
Price	−1.2339	0.0075	−1.2373	0.0072
λ	0.8903	0.0101	0.8824	0.0123
γ	0.1104	0.0062	0.1049	0.0287
G	0.3211	0.0819		
σ_ν^2	0.0015	0.0102	0.0015	0.0385
σ_ξ^2	0.0571	0.0005	0.0572	0.0006

Note: Both models include market/brand specific fixed effects.

4.2. *Demand estimates*

We now present the results for the logit demand system applied to the scanner data. In Table 4, we report the parameter estimates for the two main models with a non-linear goodwill production function and carry-over effects. The first column corresponds to the model using the parametric specification of the goodwill production function ψ, which measures the transformation of current advertising GRPs into goodwill. The second column corresponds to a model using a more flexible spline approximation of ψ. To simplify the presentation, we do not report the spline coefficients themselves. Instead, we plot the estimated spline function in Figure 3 along with the corresponding knots used (indicated by '*'). Although not reported, both specifications also include market-specific brand intercepts. Hence, the error term, ξ, only captures time-variant unobserved preference shocks for a brand. For both models, we also report the log-likelihood, and the Bayesian Information Criterion (BIC)[23] corresponding to the estimated parameters. It is evident that both models provide very similar estimates, at least for the parameters that are not related to advertising.

The signs on prices and advertising are as expected. Although not reported, all price elasticities are between −2 and −4.

For both models, we find considerable evidence of carry-over as seen in the parameter λ. This parameter captures the rate at which advertising goodwill decays over time (weeks). Suppose GRPs are purchased today such that they produce 1 unit of goodwill (recall that current advertising is converted into goodwill via the goodwill production function). At a decay rate of $\lambda = 0.89$, this single unit of goodwill today will decay to only half a unit in 6 weeks time. This finding suggests that advertising indeed exhibits strong carry-over effects into future periods and, therefore, that firms should be forward-looking when making their advertising decisions.

23 We use $BIC = (-2\log(L)) + K\log(JMN)$, where L is the partial likelihood, K is the number of parameters, J is the number of products, M is the number of markets and N is the number of weeks. This model selection criterion penalizes for the number of parameters used.

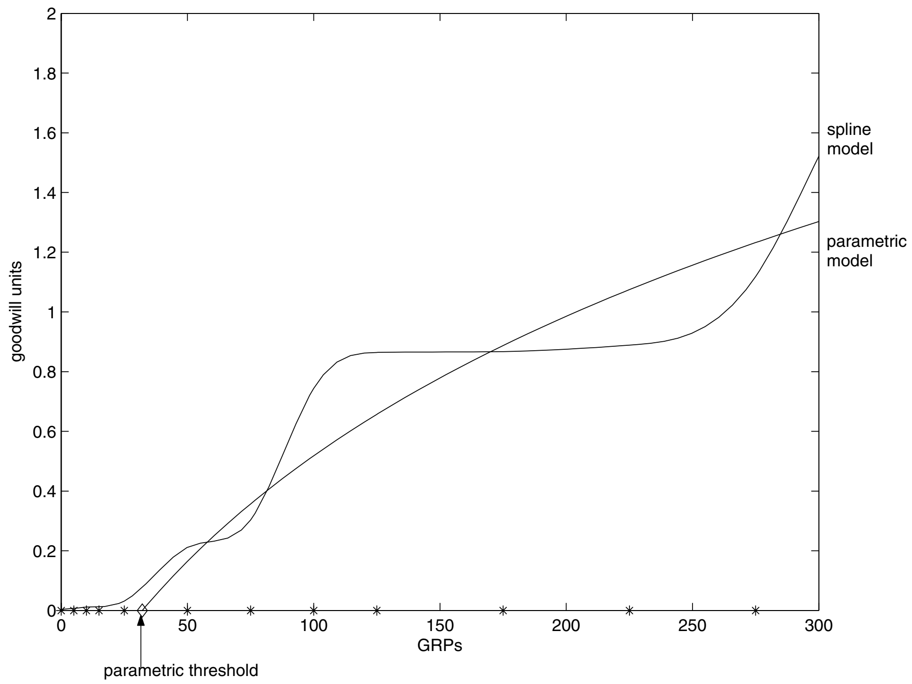

Figure 3. Estimated goodwill production function, $\psi(A)$: Parametric and spline specifications. Stars indicate location of knots used for spline approximation.

We now compare the impact of current advertising for the two models by plotting their respective goodwill production functions, $\psi(A)$, in Figure 3. The parametric estimate of the threshold effect is 32 GRPs, which rejects the hypothesis that $G = 0$ (i.e. goodwill production is not strictly concave).[24] The spline model indicates a threshold-like effect at roughly the same level of GRPs. In this respect, the parametric model appears to provide a reasonable approximation of the level below which current advertising does have no, or only little effect. For advertising levels between roughly 80 and 175 GRPs, the spline model predicts a larger advertising effect than the parametric model. In general, however, both models predict a threshold below which advertising has no effect. Most importantly, this threshold is not just an artefact of a parametric specification as it is robust to a flexible functional form specification. Finally, the spline estimates indicate that the goodwill production function passes through the origin at the threshold, $\psi(G) = 0$.[25] For the remainder of the paper, we will refer to the model with the parametric goodwill production function as our baseline model.

24 The estimate of $G = 0.3211$ in Table 4 is with respect to 100 GRPs. Therefore, the threshold is $100 \times 0.3211 = 32.11$ GRPs.

25 A potential concern with using make-goods to identify the threshold is that make-good GRPs are inherently inferior to scheduled GRPs. Hence, the threshold could have been proxying for differences in GRP quality. However, in such an instance, we would expect to observe $\psi(G) > 0$.

To assess the importance of including both carry-over effects and allowing for a non-convexity (threshold) in goodwill production, we also report estimation results from several comparison models in Table 5. By assessing in-sample fit for these comparison models to our "threshold model," we test statistically the various shape properties required for pulsing. Each of these models provides an inferior in-sample fit to our baseline model both based on the likelihood and the BIC, supporting the assumptions in our "threshold model". In model (I), we do not include advertising at all. In model (II), we include current advertising as a linear covariate in the conditional indirect utility, leading to an improved likelihood. In model (III), we also include a threshold. For the remaining models, we allow for long-run effects of advertising. In model (IV), we remove the threshold and, instead, allow for carry-over effects, which improves the fit considerably. Model (IV) has a linear goodwill production function and a linear impact of augmented goodwill, g^a, on utility. This specification gives rise to the familiar linear-distributed lag formulation of goodwill over time (e.g., Nerlove and Arrow, 1962). In model (V), we still use the linear goodwill production function, but augmented goodwill enters the conditional indirect utility in the logarithmic form, $\gamma \log(1 + g^a)$. Finally, in model (VI), we use the logarithmic goodwill production function, but do not allow for a threshold, i.e. $\psi(A) = \log(1 + A)$. Overall, we see considerable improvement in fit comparing our baseline "threshold" model to each of these alternative specifications.

In summary, our estimates indicate the presence of a threshold-like effect in the advertising response, and long-run effects of advertising on demand. In the next section, we explore the strategic implications of these findings for the firms' advertising.

4.3. Supply-side predictions

We now investigate the implications of the demand estimates for optimal (equilibrium) advertising. Most importantly, we check whether the firms in our empirical application should actually pulse in order to maximize their profits. The role of pulsing is then explored by looking at optimal advertising under some of the comparison demand specifications estimated in the previous section. To assess the importance of optimal advertising, we compare actually realized profits to the predicted profits from our advertising model. We also look at the profits of a zero-advertising regime to assess the importance of advertising.

We first calculate the equilibrium advertising policies that are implied from the demand estimates (see the appendix for the technical details of the numerical solution algorithm). The calculations are based on a discount factor of $\beta = 0.998$, which corresponds to an annual interest rate of about 11%. We checked for the uniqueness of a MPE in pure strategies by using different initial guesses of the equilibrium policies as starting values in the solution algorithm. We did not find any evidence for multiple equilibria. The calculated advertising strategies are used to simulate a series of advertising levels, market shares, and profits in each market.[26] Note that the simulated advertising levels are not in-sample predictions. In fact,

26 We simulate 10,000 weeks of the advertising game in each market. The summary statistics of the simulations are reported only for the last 9,000 weeks, in order to remove any influence of the initial conditions.

Table 5. Demand estimates of several alternative demand specifications. All models include market/brand specific fixed effects.

Model	(I) No advertising		(II) Linear advertising effect		(III) Linear advertising with threshold effect		(IV)[a] Goodwill is linear distributed lag, Γ is linear		(V)[b] Goodwill is linear distributed lag, Γ is logarithmic		(VI)[c] Full model without threshold effect	
	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Log-likelihood	−91.27		−14.34		−0.62		149.78		148.28		139.38	
BIC	686.49		840.34		898.91		598.12		610.66		628.46	
Price	−1.2353	0.0103	−1.2305	0.0102	−1.2430	0.0075	−1.2409	0.0075	−1.2399	0.0077	−1.2387	0.0077
λ							0.9158	0.0074	0.8622	0.0139	0.8642	0.0139
γ			0.0476	0.0038	0.1282	0.0197	0.0178	0.0010	0.0787	0.0062	0.0961	0.0062
G					0.0813	0.0851						
σ_ν^2									0.0169	0.4320	0.0107	0.4320
σ_ξ^2					0.0586	0.0005	0.0573	0.0005	0.0573	0.0008	0.0573	0.0008

[a] Goodwill is a linear distributed lag of advertising, and enters utility linearly, $\Gamma(g^a) = \gamma g^a$.

[b] Goodwill is a linear distributed lag of advertising, and enters utility in the logarithmic form $\Gamma(g^a) = \gamma \log(1 + g^a)$.

[c] Non-linear goodwill production function, and logarithmic effect of goodwill on utility.

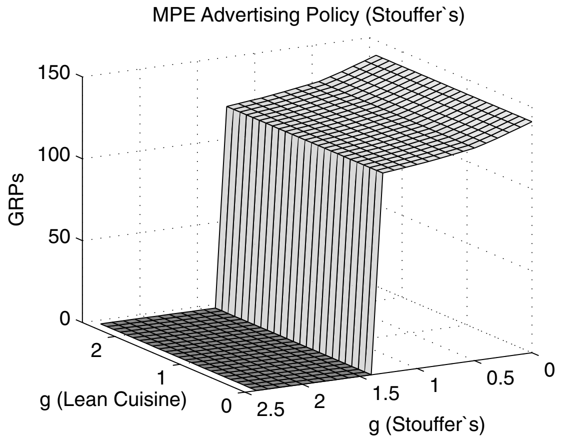

Figure 4. Optimal advertising policy.

during the estimation stage we used advertising only as a covariate, and remained agnostic as to how advertising was generated. Therefore, it is meaningless to compare a particular sequence of simulated advertising levels to the actual sequence of realized advertising levels in the data. Instead, one should compare the simulated and actual advertising *patterns*, i.e. the simulated and actual distributions of advertising.

Figure 4 shows the optimal advertising policy for the Stouffer's brand in Atlanta. The displayed policy is representative of the policies found for all brands. Advertising is decreasing in own goodwill, and has a 'kink', i.e. there is a level of goodwill at which optimal advertising abruptly drops to zero. The kink in the advertising policy is consistent with a demand-side explanation of pulsing. However, realized goodwill could always be small enough such that the firm operates on the positive part of the advertising rule. Hence, we still have to confirm pulsing through our simulations. We plot 52 weeks of simulated advertising for the Healthy Choice, Lean Cuisine, and Stouffer's brands in the left-hand column of Figure 5. It can be seen that all three brands pulse in equilibrium.[27] Table 6 provides summary statistics to compare simulated MPE advertising with the observed advertising data. Both the average level of GRPs delivered, and the average frequency of periods of no advertising are reported.[28] Average predicted advertising, i.e. the average GRPs delivered through time, is on average higher than observed advertising. This result is most pronounced for Budget Gourmet and Swanson, the two brands that advertised only little during 1991–1993. For the other brands, predicted advertising is between 15% and 48% higher than actual advertising. The difference is mainly due to a scheduling effect, as our simulated advertising patterns do not exhibit the long periods of no or only very little advertising observed in the data. Hence, our results indicate that an optimal media-schedule would involve more frequent

27 The length of the spells between weeks of positive advertising observed in Figure 5, i.e. one and two weeks, is the most common length across all brands and markets. Spells of three weeks or more occur in about 10% of all observations.

28 In a separate appendix that can be requested from the authors, we report market-specific results by brand.

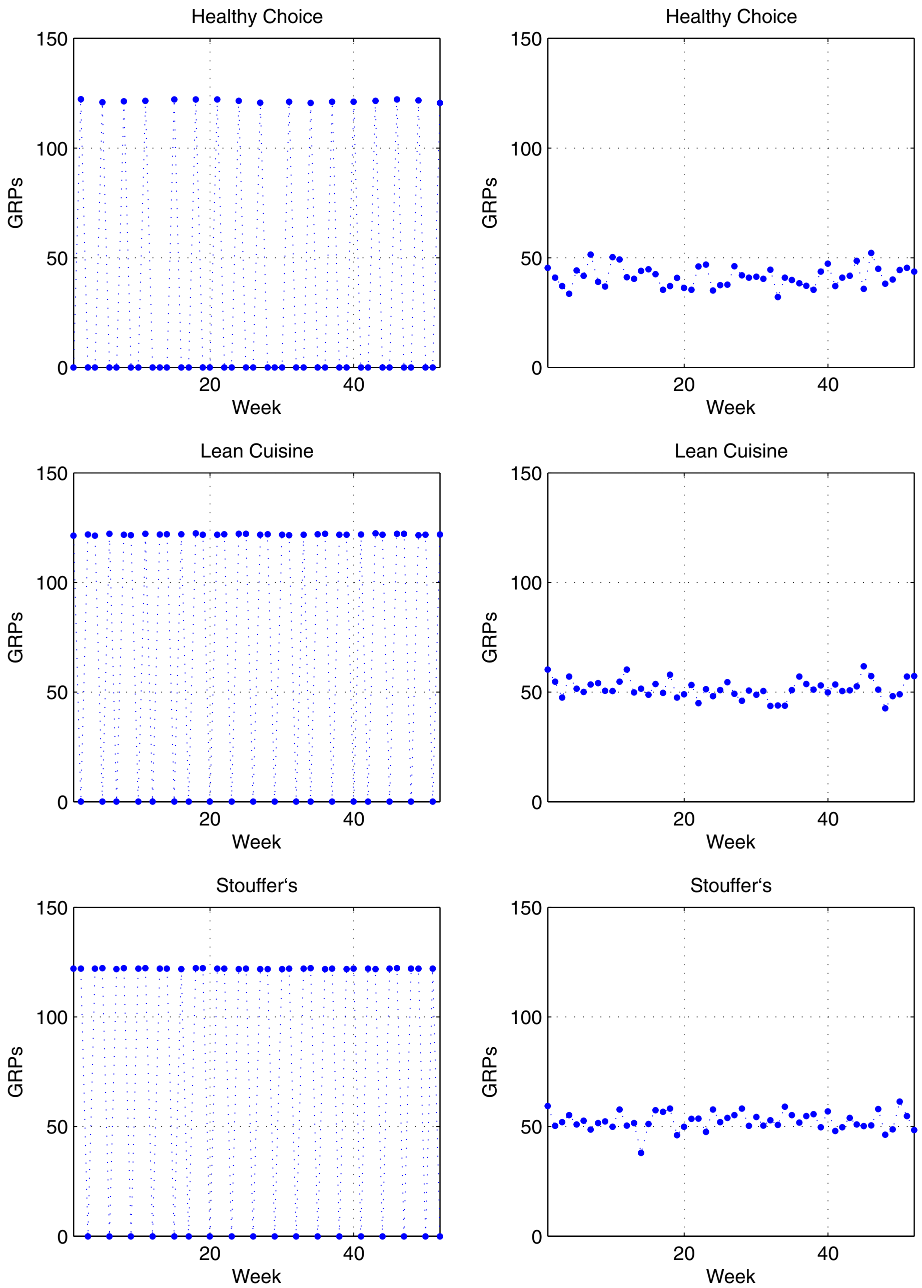

Figure 5. Simulated MPE Advertising for the Atlanta Market. The left-hand column shows the predictions of the baseline model. The right-hand side shows the results based on comparison model (V), where goodwill is a linear distributed lag of advertising.

Table 6. Advertising patterns across markets.

	Data		MPE	
Brand	Average GRPs	% weeks of 0 adv.	Average GRPs	% weeks of 0 adv.
Budget Gourmet	16.5	75.9	70.8	43.3
Healthy Choice	29.2	63.5	35.7	70.0
Lean Cuisine	34.9	57.3	40.0	66.5
Stouffer's	43.3	60.5	63.9	47.0
Swanson	9.8	87.1	54.4	55.4

Note: The table displays observed advertising, and equilibrium (five firm oligopoly) advertising under the baseline demand specification.

Table 7. Advertising patterns: Five firm oligopoly under alternative demand specifications.

	(II)[a]		(IV)[b]		(V)[d]	
Brand	Average GRPs	% weeks of 0 adv.	Average GRPs	% weeks of 0 adv.[c]	Average GRPs	% weeks of 0 adv.
Budget Gourmet	0.0	–	888.9	0.0	19.7	0.0
Healthy Choice	0.0	–	888.9	0.0	28.5	0.0
Lean Cuisine	0.0	–	1000.0	0.0	25.7	0.0
Stouffer's	0.0	–	1000.0	0.0	33.5	0.0
Swanson	0.0	–	944.4	0.0	30.2	0.0

Note: The MPE in models *(II)* and *(IV)* are restricted such that firms cannot advertise more than 1,000 GRPs per week.
[a] Linear effect of advertising on utility, no carry-over.
[b] Goodwill is a linear distributed lag of advertising, and enters utility linearly.
[c] % weeks of 0 advertising excludes observations where a specific brand never advertises in a market.
[d] Goodwill is a linear distributed lag of advertising, and enters utility in the logarithmic form.

pulses, that are also more evenly spread out through the year, than is actually observed in the data.

Next, we investigate the implications of alternative demand specifications for equilibrium advertising. Table 7 contains the summary statistics on advertising under three comparison models.[29] Model (II) is a simple demand specification, where only current advertising enters utility in a linear fashion. We see that firms do not advertise at all if advertising has no carry-over. Recall that the baseline model estimates implied a large degree of persistence in goodwill. Under the restriction of no persistence, the return on advertising does not justify its cost. Note that we obtained this result under one restriction: Firms cannot advertise more than 1,000 GRPs per week. This restriction if related to our justification of the logarithmic goodwill effect on utility that we discussed in Section 3. The logarithm ensures that the market share is strictly concave, and not S-shaped in goodwill. We found that if firms can

29 The models are numbered corresponding to the demand estimates in Table 5.

advertise large amounts in excess of 1,000 GRPs, they will in fact spend vast sums on advertising, until they reach the concave upper part of the logit share equation. Solving for these equilibria is computationally burdensome, and hence we could not verify that they even exist. If they did exist, they would be characterized by large advertising spending, that is justified by correspondingly large market shares and markups. This effect is clear from the results of model (IV), which is a specification with linear distributed lag of advertising and a linear effect of goodwill on utility. Now, firms mostly have a strong, and initially increasing incentive to advertise. Therefore, in the restricted equilibrium, most firms spend exactly 1,000 GRPs week after week.[30] In model (V), goodwill enters utility in the logarithmic form. Now, the market share equation is well-behaved, i.e. strictly concave. Goodwill is a linear distributed lag of advertising, i.e. the goodwill production function does not exhibit an S-shape or threshold. Therefore, as expected, firms do not pulse. Simulated advertising from model (V) in displayed in the right-hand column of Figure 5 . In summary, each of these comparison models fails to account for some basic patterns of the data. First, we saw that not allowing for advertising carry-over implies that firms should not advertise at all. Second, if the functional form of the advertising effect is not carefully chosen, the reverse effect can set in, and firms are predicted to advertise large amounts to capture the whole market. Third, a simple, linear shape by which advertising increases goodwill does not imply pulsing.

Villas-Boas' (1993) model predicts that firms may alternate their advertising pulses between periods. In our empirical application, we do not find evidence for this effect. The correlation between competitors' advertising levels is virtually zero. Hence, in our empirical application a firm mostly schedules its advertising to control its own goodwill and hence demand, and not to 'avoid' periods where competitors schedule their advertising.

We now focus on the profit implications of optimal advertising in a MPE. In Table 8, we report the average estimated profits for each firm corresponding to the advertising

Table 8. Profits.

		MPE		Zero adv.		Even adv.	
Brand	Data total profits	Total profits	% diff. wrt. actual	Total profits	% diff. wrt. actual	Total Profits	% diff. wrt. MPE
Budget Gourmet	62.47	65.12	4.25	61.39	−1.73	63.89	−1.90
Healthy Choice	41.74	42.37	1.51	41.53	−0.51	40.11	−5.33
Lean Cuisine	47.24	47.73	1.04	46.56	−1.44	44.94	−5.86
Stouffer's	66.54	67.46	1.37	63.88	−4.01	66.21	−1.85
Swanson	55.38	58.12	4.95	55.57	0.35	56.02	−3.61

Note: Total profits are reported in million dollars per year in all 18 markets. "Even" Advertising refers to a policy whereby firms reallocate the total number of predicted GRP's under MPE advertising, and schedule advertising evenly week after week.

30 Average GRPs delived as reported in Table 7 are sometimes less than 1,000 GRPs because in a small number of markets, some of the firms do not advertise at all under the imposed maximum GRP restriction.

levels observed in the data, and the advertising predicted by the MPE. We also report the average profit estimates if firms never advertised. The MPE advertising policies lead to moderate improvements over the observed profit levels. Interestingly, the two firms that are observed to advertise the least, Swanson and Budget Gourmet, are predicted to gain the most from switching to the MPE advertising policies. The results suggest that these two firms deviate the most from their optimal advertising levels; in particular, they advertise too infrequently. Next, we examine the zero-advertising regime. With the exception of Swanson, if firms did not advertise at all, they would incur moderate profit losses relative to the observed advertising regime. Hence, firms are overall better off with the ability to advertise as the increase in category demand outweighs the potential "prisoner's dilemma" effect of advertising.[31]

Finally, we examine the implications of scheduling by comparing the profitability of pulsing to a continuity strategy whereby firms advertise every period. To simulate continuity advertising, we take the total advertising budget predicted by the MPE advertising policy, but schedule advertising evenly across weeks, and not in a pulsing fashion. Hence, firms spend the same amount on advertising under both scenarios and only the scheduling of the ads differs. Our results confirm the importance of the pulsing aspect of advertising. Compared to MPE advertising, under continuity strategies firms would incur profit losses that range between 1.9 and 5.9 percent.[32] Two brands, Healthy Choice and Lean Cuisine, would be even better off without advertising at all compared to the sub-optimal, continuity advertising policy.

To summarize, we find that optimal advertising does indeed involve pulsing as an advertising scheduling strategy. Recall once again that our demand estimates were obtained without imposing any specific supply-side assumptions on the estimation procedure. Hence, we did not make any assumptions on what forms or levels of advertising were optimal when we estimated demand. Therefore, pulsing is a consequence of our demand estimates, and the additional assumption that firms behave optimally to maximize their expected PDV of profits. Previous research has shown that pulsing can be optimal in theory, and has empirically documented that in some markets, advertising response curves exhibit an S-shape, or a related non-linear form. Our study, however, is the first to show that based on an estimated demand system, pulsing does arise as the outcome of optimizing behavior, accounting for oligopolistic competition among firms.[33]

4.4. Pulsing under oligopoly and monopoly

We now examine how predicted advertising changes under an alternative, monopolistic market structure, where one firm owns all brands. We found that it is very time intensive

31 As before, market-specific profit measures are reported in a separate appendix.

32 The average profit loss would be \$1.93 million, on average, in absolute terms.

33 Simon (1982) shows that a particular form of pulsing, i.e. the alternation between 'high' and 'low' advertising levels, is implied by his demand estimates. However, as discussed in the introduction, his demand system is highly restrictive, and he neither allows for competition in the demand system nor for competitive interaction among firms.

Table 9. Oligopoly vs. monopoly advertising.

Brand	3 Firm oligopoly			3 Product monopoly		
	Average GRP's	% weeks of 0 adv.	Average GRP's > 0	Average GRP's	% weeks of 0 adv.	Average GRP's > 0
Budget Gourmet	70.5	38.8	107.6	64.9	42.9	105.7
Stouffer's	63.0	42.0	105.0	57.9	46.7	104.5
Swanson	55.7	49.2	105.4	50.9	52.9	103.8

to solve for the optimal advertising policy of a multi-product monopolist.[34] Therefore, we restrict the analysis to a simplified market, where only the three brands that advertise most in the five firm MPE (Budget Gourmet, Stouffer's, and Swanson) are sold. Table 9 compares the advertising outcomes under oligopoly and multi-product monopoly. Under monopoly, the firm internalizes the business stealing effect from increased advertising for a specific brand. Therefore, average advertising decreases by about 8% relative to oligopoly. The prediction of pulsing, however, is unaffected by the particular market structure. The reduction in average advertising is achieved through changes both along the intensive margin, where the average level of advertising in non-zero advertising periods is reduced, and along the extensive margin, where the fraction of periods without advertising is increased.

These findings provide evidence that the degree of competition does influence qualitative aspects of how firms pulse. However, the incidence of pulsing seems to be driven by the manner in which advertising impacts demand, and not by the nature of competition.

4.5. Discussion

Our model predicts pulsing, and thus matches some of the observed advertising patterns. On the other hand, some data features are not accounted for. In particular, actual advertising in the Frozen Entrée category exhibits some long 'blocks' of little or no advertising, and the levels of positive advertising are much more variable than the predicated advertising levels from our model. The purpose of our paper is to characterize the consequences of the shape of the advertising response, and competition, for optimal advertising scheduling. We remained agnostic as to whether observed advertising is actually optimal while formulating and estimating product demand, and solving for the equilibrium strategies. An alternative approach would be to *assume* that firms already advertise optimally, and then draw inferences about optimal advertising from the data. Note that the demand estimator under this approach will generally be biased if the assumptions on competitive conduct are incorrect. Furthermore, although our discussions with industry participants revealed that firms understand the basic demand-side rationale for pulsing, we did not

34 Specifically, the maximization on the right-hand side of the Bellman equation (7) has to be carried out with respect to a multi-dimensional control variable, which substantially increases the computational burden relative to the one-dimensional case.

find any evidence that firms know exactly how to solve for a fully optimal advertising schedule, which is based on a complicated dynamic optimization problem. Therefore, our interpretation of the data is that the long blocks of zero advertising are due to binding advertising constraints. This finding suggests that firms might be better off by increasing their advertising budgets. The higher variability of positive advertising levels in the data, on the other hand, may be due to factors that we do not model. First, the effectiveness of advertising across firms may be different, i.e. γ may be brand/firm specific. Second, advertising costs vary through time, while we simulate optimal advertising at the average observed GRP cost. Third, contracted GRP levels, and actually realized GRP levels are generally not the same. Our model predicts only the ex-ante, contracted GRP levels. Allowing realized GRP levels to differ from the ex-ante levels would add noise to predicted advertising, which would trivially improve the similarity between the volatilities of simulated and observed advertising. Fourth, we do not model the institutional practice of make-goods, which would add additional noise to the simulated advertising series.

5. Conclusions

Using data from the Frozen Entrée category, we find that optimal advertising scheduling exhibits pulsing. This result is driven by the demand side, in particular the presence of a threshold effect in the response of goodwill to advertising. Our advertising predictions allow for strategic interaction among firms, and are based on the Markov perfect equilibrium concept that we use to solve for the outcome of the advertising game. However, while competition influences the level and frequency of optimal advertising, we find that pulsing as the optimal form of advertising scheduling arises under both competition and monopoly. Furthermore, the advertising schedules for different brands are largely uncorrelated. Our estimates reveal that demand has a non-linear response to advertising, including a threshold effect, and that the carry-over effect of current advertising into future periods is statistically and economically significant. Alternative demand specifications with simpler advertising response functions do not fit the data as well as the baseline model and do not yield pulsing as an optimal advertising strategy. An important aspect of this analysis is that we do not impose any supply-side restrictions during the demand estimation stage. The data used contain unusually good advertising measures, that are based on gross rating points, and not dollar expenditures.

We also find that firms would increase their profits if they adopted our recommended optimal advertising schedules. While the gains appear moderate when measured in percentage terms, they are between 1.3 and 7 million dollars per year in absolute terms.[35] At an 11% interest rate, the present value of these improvements is between 11 and 63 million dollars. Similarly, we find that if firms simply stopped advertising entirely, they would, on average, lose 2.3 million dollars in profits per year. Also, if firms reallocated their optimal advertising budgets and spread advertising evenly throughout the year, profits would decrease by about 1.9 million dollars per year, on average, relative to optimal advertising. That is, a continuity

35 These numbers are based on a projection of the profits in our 18 sample markets to the whole U.S.

strategy is less profitable than a pulsing strategy and, hence, scheduling does emerge as an important aspect of advertising.

To estimate the parameters on the demand-side, we use a partial maximum likelihood (PML) procedure. The nonlinearities of the advertising response function make it infeasible to use simple non-linear least squares methods. The PML procedure is a relatively straightforward method for obtaining consistent demand estimates. We also provide a standard-error correction to account for the (potential) time-dependence in our data.

On the supply side, we use numerical dynamic programming techniques that provide considerable flexibility in the types of demand models we can consider. Unlike the extant theoretical literature in this area, we use a logit demand system, which is more 'realistic' in that it is consistent with the types of models typically used to estimate demand with consumer data. We solve for a Markov perfect equilibrium that involves all five firms that advertise in our market. Previous research has typically ignored competition or focused on duopoly models. As discussed above, competition does affect the qualitative aspects of pulsing.

As stated previously, an advantage of our computational approach is that it can accommodate a wide variety of demand systems. For our current question regarding pulsing, we have focused on the simple homogeneous logit demand system. In principle, one could introduce heterogeneity into the estimated demand system. However, it is unlikely that omitted heterogeneity bias would have a first order impact on our findings regarding pulsing. In fact, the usual result is that omitting heterogeneity understates the mean response to marketing variables. Thus, controlling for heterogeneity would likely increase the mean response to advertising which would merely strengthen our current findings. Studying the role of heterogeneity in the presence of dynamic advertising is left as an interesting area for future research.

An important issue that is currently not addressed is the potential for bias due to price and advertising endogeneity. A resolution of these issues is difficult in our estimation framework. However, the exogenous variation in advertising due to the practice of make-goods allows us to identify the threshold in the goodwill production function, and hence the main conclusion of our paper–firms should optimally pulse–is robust to the possibility that some other parameter estimates are biased. Furthermore, the inclusion of brand-specific fixed-effects that are allowed to vary across markets mitigates the potential for bias due to price endogeneity. Nonetheless, a more satisfactory treatment of econometric endogeneity in our estimation framework is an important area for future research.

Finally, due to the aggregate nature of our data, we are not able to provide a more microeconomic rationale for the role of advertising. By using a flexible functional form for the advertising response, we have tried to be agnostic about the role of advertising on demand. Nevertheless, future research would clearly benefit from data that permitted addressing the role of copy and creative on the effect of advertising, and the exact way in which advertising influences household demand.[36] This type of data would help control for differences across firms in the marginal impact of advertising on demand.

36 Ackerberg (2001), for example, shows how to estimate and distinguish between prestige and informative advertising effects.

Appendix A: Pulsing and the shape of the goodwill production function

We stated in Section 3 that in a monopoly, pulsing will not arise if the goodwill production function is globally concave. The proof of this statement is simple. Consider the goodwill production function $g^a = \chi(g, A)$. χ is increasing in its arguments, globally concave in advertising at any g, and twice continuously differentiable. Furthermore, we assume that the marginal return to advertising is not increasing in the amount of already accumulated goodwill,

$$\frac{\partial^2 \chi(g, A)}{\partial g \partial A} \leq 0. \tag{16}$$

The goodwill production function in our model, $g^a = g + \psi(A)$, is a special case of this technology. Note that in our model the transition probability of the state vector, $p(\cdot|g, A)$, is only conditional on augmented goodwill, and not separately on both goodwill and advertising. Therefore, the transition probability can be written as $p(\cdot|g_j^a)$. Similarly, the current period profit function, π, is only dependent on augmented goodwill. Let $\hat{\pi}$ be the current profit flow, gross of advertising, such that $\pi = \hat{\pi} - kA$. Assume that both $\hat{\pi}(g^a)$ and the expected future value, $\int V(g')p(g'|g^a)dg'$, are increasing, strictly concave, and twice continuously differentiable in augmented goodwill. Let A^* be the optimal advertising level at the goodwill stock g, and $g^{a*} = \chi(g, A^*)$ be the corresponding augmented goodwill stock. If $A^* > 0$, this augmented goodwill level satisfies the first order condition

$$\frac{\partial \hat{\pi}(g^{a*})}{\partial g^a} \frac{\partial \chi(g^{a*}, A^*)}{\partial A} - k + \beta \int V(g') \frac{\partial p(g'|g^{a*})}{\partial g^a} \frac{\partial \chi(g^{a*}, A^*)}{\partial A} dg' = 0. \tag{17}$$

Now consider another goodwill level $g_0 < g^{a*}$. By assumption, χ is concave in A, and the profit function and expected future value are strictly concave. From assumption (16) and concavity, we have

$$\frac{\partial \chi(g^{a*}, A^*)}{\partial A} \leq \frac{\partial \chi(g^{a*}, 0)}{\partial A} \leq \frac{\partial \chi(g_0, 0)}{\partial A}.$$

Hence, as both the profit and value functions are strictly concave,

$$\begin{aligned} k &= \frac{\partial \chi(g^{a*}, A^*)}{\partial A} \left(\frac{\partial \hat{\pi}(g^{a*})}{\partial g^a} + \beta \int V(g') \frac{\partial p(g'|g^{a*})}{\partial g^a} dg' \right) \\ &< \frac{\partial \chi(g_0, 0)}{\partial A} \left(\frac{\partial \hat{\pi}(g_0)}{\partial g^a} + \beta \int V(g') \frac{\partial p(g'|g_0)}{\partial g^a} dg' \right). \end{aligned}$$

Therefore, $A = 0$ is not the optimal advertising level at the goodwill stock g_0. We conclude that if the firm advertises at some state g, thus producing the augmented goodwill level g^{a*}, it will also advertise in the next period as long as goodwill depreciates between periods, $g_0 < g^{a*}$. Therefore, pulsing cannot arise as a profit-maximizing strategy.[37]

Appendix B: Computational details

We solve for the MPE of the dynamic advertising game using a simple adaptation of a policy iteration algorithm to the case of multiple decision makers.[38] This algorithm takes some initial guess of the strategy profile, $\sigma^0 = (\sigma_1^0, \dots, \sigma_J^0)$, and then proceeds according to the following steps:

1. For the strategy profile σ^n, calculate the corresponding value functions V_j^n for each of the J firms. These value function are defined by the Bellman equation (7), where the maximization on the right hand side is not actually carried out, but instead the current guess of the strategy profile σ^n is used.
2. If $n > 0$, check whether the value functions and policies satisfy the convergence criteria, $\|V_j^n - V_j^{n-1}\| < \varepsilon_V$ and $\|\sigma_j^n - \sigma_j^{n-1}\| < \varepsilon_\sigma$. If so, stop.
3. Update each firm's strategy from the Bellman equation (7). In contrast to step 1, the maximization on the right hand side is now carried out. Denote the resulting new policies and value functions by σ^{n+1} and V^{n+1}, and return to step 1.

The main remaining computational details concern the approximation of the value functions and policies, and the calculation of the integral in the Bellman equation. The value functions are represented on a grid $\mathcal{G} = \{g^i | i = 1, \dots, G\}$. This grid is constructed by first discretizing each goodwill axis into the points $g_{j1} < g_{j2} < \cdots < g_{jK}$, and then collecting all $G = K^J$ resulting J dimensional points. Outside $\mathcal{G}$, the value function is evaluated through bilinear interpolation. The numerical integrals are calculated by Gauss-Hermite quadrature, and by Monte Carlo integration.

We note that function approximation on a grid can be computationally expensive. An usually faster alternative is polynomial approximation (Benítez-Silva et al., 2003). However, we found that this method did not perform well in our specific application, due to the shape of the equilibrium advertising policies. These policies contain a kink, i.e. a discontinuity, which is hard to approximate through a polynomial. Consequently, we found that the policies from a polynomial approximation were often significantly different from the results based on a very fine grid.

37 In our model, goodwill today could be at least as large as augmented goodwill in the last period, $g_t \geq g_{t-1}^a$, due to the depreciation shock ν_t. Then, advertising in period t would in fact be zero. However, from the estimated standard error of ν and the regular predicted pulsing patterns in Figure 5 we can immediately rule out that the possible occurrence of 'favorable' depreciation shocks drives the predicted pulsing behavior in our empirical application.

38 Rust (1995) and Judd (1998) provide overviews of dynamic programming techniques.

Acknowledgments

The data for this study were made available by Dennis Bender. The authors benefitted from discussions with Mike Duffy about media-planning. We are grateful to Pradeep Chintagunta, Tim Conley, Gary Erickson, Robert McMillan, Demetrios Vakratsas, Naufel Vilcassim, Miguel Villas-Boas, and Peter Rossi for comments and suggestions. The authors also benefited from the comments of seminar participants at the Columbia GSB, the Department of Justice, the Federal Trade Commission, the London Business School, the Ohio State University, the University of Arizona, the University of British Columbia, UC Berkeley, UCLA, the University of Toronto, Washington University in St. Louis, the joint workshop between the University of Waterloo, Wilfried Laurier University and CIGI, the 2002 Marketing Science Conference in Edmonton, the 2003 Summer Institute in Competitive Strategy at Berkeley, and the 2003 Winter Marketing Camp at Leuven. We would like to acknowledge the Kilts Center for Marketing at the University of Chicago for providing research funds. The first two authors were also supported by a True North Communications Inc. Scholarship. The third author was also supported by the Beatrice Foods Company Faculty Research Fund Scholarship.

References

Ackerberg, D.A. (2001). "Empirically Distinguishing Informative and Prestige Effects of Advertising." RAND Journal of Economics 32(2), 316–333.

Assmus, G., J.U. Farley, and D.R. Lehmann. (1984). "How Advertising Affects Sales: Meta-Analysis of Econometric Results." Journal of Marketing Research 21, 65–74.

Baron, R. (2003). "Planning's Top 5 Questions." TelevisionWeek 22(12), 23–32.

Benítez-Silva, H., G. Hall, G.J. Hitsch, G. Pauletto, and J. Rust. (2003). "A Comparison of Discrete and Parametric Approximation Methods for Continuous-State Dynamic Programming Problems." manuscript.

Benkard, L. (2001). "A Dynamic Analysis of the Market for Wide-bodied Commercial Aircraft." Review of Economic Studies (forthcoming).

Berry, S. (1994). "Estimating Discrete-Choice Models of Product Differentiation." Rand Journal of Economics 25, 242–62.

Berry, S., J. Levinsohn, and A. Pakes. (1995). "Automobile Prices in Market Equilibrium." Econometrica 63(4), 841–890.

Bronnenberg, B. (1998). "Advertising Frequency Decisions in a Discrete Markov Process Under a Budget Constraint." Journal of Marketing Research 35, 399–406.

Cannon, H.M., J.D. Leckenby, and A. Abernethy. (2002). "Beyond Effective Frequency: Evaluating Media Schedules Using Frequency Value Planning." Journal of Advertising Research 42(6), 33–47.

Chen, X. and T.G. Conley. (2001). "A New Semiparametric Spatial Model for Panel Time Series." Journal of Econometrics 105, 59–83.

Chintagunta, P.K., V. Kadiyali, and N.J. Vilcassim. (1999). "Investigating Dynamic Multi-firm Market Interactions in Price and Advertising", Management Science 45(4), 499–518.

Chintagunta, P.K., J. Dube, and K. Yong Goh. (2005). "Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household-Level Brand Choice Models." Management Science 51(2).

Clarke, D.G. (1976). "Econometric Measurement of the Duration of Advertising Effect on Sales." Journal of Marketing Research 13, 345–357.

Doraszelski, U. and S. Markovich. (2003). "Advertising Dynamics and Competitive Advantage." Working Paper, Hoover Institution.

Doraszelski, U. and M. Satterthwaite. (2003). "Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity." manuscript.

Ephron, E. (2002). "Median Scheduling and Carry-over Effects: Is Adstock a Useful TV Planning Tool?" Journal of Advertising Research 42(4), 66–71.

Erdem, T. and M.P. Keane. (1996). "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." Marketing Science 15, 1–20.

Erickson, G. (1995). "Advertising Strategies in a Dynamic Oligopoly." Journal of Marketing Research 32, 233–237.

Erickson, G. (2002). Dynamic Models of Advertising Competition (International Series in Quantitative Marketing, 13), 2nd edition. Kluwer Academic Publishers.

Ericson, R. and A. Pakes. (1995). "Markov Perfect Industry Dynamics: A Framework for Empirical Work." Review of Economic Studies 62(1), 53–82

Feinberg, F.M. (1992). "Pulsing Policies for Aggregate Advertising Models." Marketing Science 11(3), 221–234.

Hitsch, G.J. (2004). "An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty." Marketing Science (forthcoming).

Jones, J.P. (1995). "Single-Source Research Begins to Fulfill its Promise." Journal of Advertising Research 35(3), 9–16.

Judd, K.L. (1998). Numerical Methods in Economics. MIT Press.

Kingman, M. (1977). "Admen See Pulsing as Way to Beat Soaring TV Time Costs." Advertising Age 48(27), 27–29.

Krugman, H.E. (1972). "Why Three Exposures May be Enough." Journal of Advertising Research 12, 11–14.

Leckenby, J.D. and J.D. Kim. (1994). "How Media Directors View Reach/Frequency Estimation: Now and a Decade Ago." Journal of Advertising Research 34(5), 9–21.

Leeflang, P.S.H. and D.R. Wittink. (1992). "Diagnosing Competitive Reactions Using (Aggregated) Scanner Data." International Journal of Research in Marketing 9, 39–57.

Leeflang, P.S.H. and D.R. Wittink. (1996). "Competitive Reaction Versus Consumer Response: Do Managers Overreact?" International Journal of Research in Marketing 13, 103–119.

Lilien, G.L., P. Kotler, and K.S. Moorthy. (1992). Marketing Models. New Jersey: Prentice Hall.

Mahajan, V. and E. Muller. (1986). "Advertising Pulsing Policies for Generating Awareness for New Products." Marketing Science 5(2), 110–111.

Maskin, E. and J. Tirole. (2001). "Markov Perfect Equilibrium." Journal of Economic Theory 100(2), 191–219.

Montgomery, D.B., M.C. Moore, and J.E. Urbany. (2005). "Reasoning About Competitive Reactions." Marketing Science 24(1), 138–149.

Naik, P.A., M.K. Mantrala, and A.G. Sawyer. (1998). "Planning Media Schedules in the Presence of Dynamic Advertising Quality." Marketing Science 17, 214–235.

Naples, M.J. (1979). Effective Frequency. New York: NY Association of National Advertisers.

Nerlove, M. and K.J. Arrow. (1962). "Optimal Advertising Policy Under Dynamic Conditions." Economica 29, 129–142.

Nevo, A. (2001). "Measuring Market Power in the Ready-To-Eat Cereal Industry." Econometrica 69(2), 307–342.

Pakes, A. and P. McGuire. (1994). "Computing Markov Perfect Nash Equilibrium: Numerical Implications of a Dynamic Differentiated Product Model." Rand Journal of Economics 25(4), 555–589.

Rao, A.G. and P.B. Miller. (1975). "Advertising/Sales Response Functions." Journal of Advertising Research 15, 7–15.

Rust, J. (1996). "Numerical Dynamic Programming in Economics." In H.M. Amman, D.A. Kendrick, and J. Rust (eds.), Handbook of Computational Economics, Elsevier Science B.V., pp. 619–729.

Samuelson, L. and M.J. Roberts. (1988). "An Empirical Analysis of Dynamic Non-Price Competition in an Oligopolistic Industry." Rand Journal of Economics 19, 200–219.

Sasiensi, M.W. (1971). "Optimal Advertising Expenditures." Management Science 18(4), 64–72.

Shugan, S. and S. Radas. (2003). "Selective Marketing Strategies: Implications of Response Thresholds and Capacity Constraints." Working Paper, University of Florida.

Simon, J.L. (1969). "New Evidence for No Effect of Scale in Advertising." Journal of Advertising Research 9, 38–41.

Simon, H. (1982). "ADPULS: An Advertising Model with Wearout and Pulsation." Journal of Marketing Research 19, 352–363.

Slade, M.E. (1995). "Product Rivalry with Multiple Strategic Weapons: An Analysis of Price and Advertising Competition." Journal of Economics and Management Strategy 4(3), 445–476.

Vakratsas, D., F.M. Feinberg, F.M. Bass, and G. Kalyanaram. (2004). "The Shape Of Advertising Response Functions Revisited: A Model of Dynamic Probabilistic Thresholds." Marketing Science 23(1), 109–119.

Villas-Boas, M.J. (1993). "Predicting Advertising Pulsing Policies in an Oligopoly: A Model and Empirical Test." Marketing Science 12, 88–102

Villas-Boas, M. and R. Winer. (1999). "Endogeneity in Brand Choice Models." Management Science 45(10), 1324–1338.

White, H. (1982). "Maximum Likelihood Estimation of Misspecified Models." Econometrica 50(1), 1–25.

Wittink, D.R. (1979). "Exploring Territorial Differences in the Relationship between Marketing Variables." Journal of Marketing Research 15, 145–155.

Wooldridge, J.M. (2001). Econometric Analysis of Cross Section and Panel Data. Boston: MIT Press.