Semi–Structural Models of Advertising Competition*

Joris Pinkse
Department of Economics
The Pennsylvania State University
joris@psu.edu

Margaret Slade
Department of Economics
University of Warwick
m.slade@warwick.ac.uk

March 2005

ABSTRACT
We propose a semi–structural discrete–choice model that can be used to estimate static or dynamic decision rules. By semi–structural we mean that our estimator is motivated by a structural model, and we estimate the choice rules that are associated with that model. However, we do not uncover the underlying structural parameters. Our estimator can be used to advantage when games involve a rich set of choices that might be correlated across decision makers and when the object is to uncover the effect of historic behavior (on, for example, prices, sales, and profits). We apply our discrete–choice estimator to study decision rules for a dynamic game of price and advertising competition.

Journal of Economic Literature classification numbers: C51, L13, L81

Keywords: Dynamic discrete choice, Spatial dependence, Price and advertising competition

*We would like to thank participants at the following seminars: the EC^2 meeting in Marseille, the Universiy of Warwick IO Workshop, Coventry, and the CEP/IFS Productivity and IO Workshop, London.

1 Introduction

In this paper, we propose a semi–structural discrete–choice model that can be used to estimate static or dynamic decision rules. By semi–structural we mean that our estimator is motivated by a structural model, and we estimate the choice rules that are associated with that model. However, we do not uncover the underlying structural parameters. We apply our discrete–choice estimator to study decision rules for a dynamic game of price and advertising competition without imposing the restrictions that are implied by a particular equilibrium of the game.

A number of practical problems are encountered when attempting to estimate static discrete–choice models. First, due to the nonlinear nature of the conditional expectation function (CEF), it can be difficult to allow for endogenous regressors. When the endogenous regressors are discrete, moreover, multiple equilibria are common; see Tamer (2003) for a discussion. Second, the nonlinear CEF also complicates the inclusion of fixed effects when the number of time periods is small. This difficulty leads many researchers to choose a simple conditional–logit model. However, there are a number of drawbacks that are associated with that model. Third, most discrete–choice estimators are inconsistent when errors are heteroskedastic. If the heteroskedastic form is known, parametric estimators can be constructed. If it is unknown, however, the coefficient estimators generally converge at a slower rate (see e.g. Manski, 1975, and Chamberlain, 1992). Finally, with most discrete–choice estimators, it is difficult to correct for spatial dependence (i.e., cross–sectional error correlation) of a general form, particularly in the presence of endogenous right hand side variables.

Additional complications arise when estimating a structural dynamic discrete–choice model. For example, it is difficult to handle a rich set of discrete and continuous choices or a large (discrete–valued) state vector in a structural context. In addition, it is usually necessary to make strong distributional assumptions concerning the errors.

To illustrate the second point, most researchers who estimate dynamic discrete–choice games assume that players possess private information.[1] Moreover, they assume that private information is i.i.d. across players (independent private values). In many applications, however, it seems natural to assume that private information is correlated. Correlation can arise, for example, when the same firm (but not the same decision maker) is involved in many markets or when retail outlets are in close geographic proximity to one another.

Our estimator is simple compared to a full structural analysis, but can handle several of the above complications in a straight–forward fashion. Our application involves a repeated, multi–stage game. Specifically, in every period decisions are made

[1] See, e.g., Seim (2004), Aguirregabiria and Mira (2004), Bajari, Benkard, and Levin (2004), Pakes, Ostrovsky, and Berry (2004), and Pesendorfer and Schmidt–Dengler (2004). Seim's mode is static, and the others are dynamic.

that can be either binary or continuous. Binary decision are modelled as functions of covariates and rivals' imputed probabilities of making certain choices, not of their actual choices. This setup is both a better fit for the application and, under mild conditions, it removes the possibility of multiple equilibria of the period–t game. Furthermore, we allow for spatial and time series dependence of a fairly general nature.

The simplicity of our methodology, however, is not obtained costlessly. In particular, semi–structural models are less suitable for policy analysis. As Lucas (1976) pointed out several decades ago, only the underlying structural parameters are invariant to policy changes. When static problems involve changes in market structure due to, for example, mergers or the introduction of new products, our technique will not be appropriate. Furthermore, much of the dynamic discrete–choice literature has focused on uncovering the parameters of an entry or other adjustment–cost function, and that exercise also requires a structural model.

Since there are costs and benefits associated with both sorts of models, the one that is chosen should depend on the application. We believe that our estimator can be used to advantage when games involve a rich set of choices that might be correlated across decision makers and when the object is to uncover the effect of historic behavior (on, for example, prices, sales, and profits). If one is interested in evaluating changes in market structure due to, for example, entry or the introduction of superior technologies, then our estimator is less useful.

We use our estimator to recover the choice rules for a dynamic game of price and advertising competition. The players in that game are manufacturers of brands of a differentiated product that is sold in grocery stores (saltine crackers in the application) who choose whether to advertise the brand, whether to feature it in an aisle display, and what price to charge. The first two decisions are discrete, whereas the third is continuous. Furthermore, since the effects of advertising persist for more than one period, the model is dynamic. We use the estimated model to assess the static and dynamic consequences of advertising for own and rival prices and sales. There are, however, many other possible applications, both static and dynamic. Indeed, our estimator can be used whenever choices are discrete.

The organization of the paper is as follows. The next section discusses some related literature. Section 3 specifies the game and section 4 presents the econometric model. The application follows. In particular, section 5 introduces the market and the data, section 6 presents the econometric results and the comparative–dynamic exercises, and section 7 concludes.

2 Related Literature

Applications

Recently, a number of researchers have estimated structural models of the effects of advertising. Some have been static [Goeree 2004 (personal computers) and Shum 2004 (breakfast cereals)] and others have been dynamic [Erdem and Keane 1996 (detergents), Anand and Shachar 2001 (network TV programs), and Ackerberg 2003 (yogurt)]. Almost all, however, have looked at how advertising changes consumers' brand choices. As a consequence, most have used household data to estimate models of consumer demand.[2] We, in contrast, are interested in modeling how advertising changes firms' decisions. For that reason, we use brand–level data and estimate firms' dynamic decision rules. In other words, whereas in most previous work the discrete decision is a consumer's choice of brand to purchase, in our work the discrete decision is a firm's choice of whether to advertise. Furthermore, we are interested in assessing how advertising changes firms' future strategic interactions.

Our work is closest to that of Slade (1995), who assesses firm price and advertising strategies. In contrast to that work, however, we are able to handle the endogeneity problem in a more satisfactory fashion and to model cross–sectional error correlation. Furthermore, the questions that we ask are more truly dynamic. For example, Slade (1995) looks at whether advertising expands the market as a whole or whether it merely shifts market shares. However, only the direct effect of past advertising on current demand is calculated. We, in contrast, also assess the indirect effect that occurs because own advertising changes rival future choices.

There is also a growing literature on modeling structural dynamic discrete–choice games. Almost all of the applications assess the entry decision in a panel of markets through either econometric estimation (Aguirregabiria and Mira (2004) and Pesendorfer and Schmidt–Dengler (2004)) or Monte Carlo simulations (Bajari, Benkard, and Levin (2004) and Pakes, Ostrovsky, and Berry (2004)).

Within that literature, our work is closest to that of Bajari, Benkard, and Levin (2004), who perform a two–stage estimation. In the first stage, they estimate the firms' decision rules, and in the second, they make use of a set of inequalities that are implied by the equilibrium of the game to estimate the structural parameters. We estimate the decision rules, imposing more structure on those rules than they do, but do not impose the equilibrium conditions or uncover the structural parameters.

Spatial, Discrete–Choice Estimators

Our estimator is based on earlier work by Pinkse, Shen, and Slade (2004).[3] In

[2] Goeree (2004) uses aggregate brand–level data.

[3] For an overview of our research program, see Slade (2004). For applications involving continuous

that paper, we provide a new central limit theorem for spatial processes under weak conditions which, unlike previous results, are plausible for most economic applications, and we show that the one–step ('continuous updating') GMM estimator is consistent and asymptotically normal under weak conditions that allow for generic spatial and time series dependence. We use our procedure to estimate a dynamic spatial probit with fixed effects that enables us to study operating decisions for mines in a real–options context. That application, however, is not strategic, and decisions are not made jointly.

In this paper, we have a latent–variable equation of the form

$$y_b^* = E\Big(\alpha \sum_{b' \neq b} y_{b'} + \beta' x_b + u_b | \mathcal{I}_b\Big), \tag{1}$$

where y_b is the observed choice of firm (brand) b, $\mathcal{I}_b$ is the information available to firm b prior to making its decision, and where the u_b's are error terms which can be dependent across firms. In other words, a firm's desired choice depends on the actual choices of its rivals. However, rival choices are not observed, which means that firms base their decisions on expected values of y. The two models differ in other respects as well. For our purposes, however, the most important is that choices are made jointly and strategically here. Nevertheless, this research draws heavily on the results from our earlier paper.

3 The Game

This section describes the players, the game that they play, and the strategies that they use. We are interested in a dynamic game of price and advertising competition among brands of retail products. However, the game that we discuss could be used as a basis for estimating other sorts of dynamic strategic interactions.

We assume that there are B players (firms or brands), $b = 1, \ldots, B$,[4] that are engaged in oligopolistic rivalry, and T periods, $t = 1, \ldots, T$, with $T \leq \infty$.[5] There can also be many markets, $s = 1, \ldots, S$, but for expositional clarity we suppress market notation. The prevailing conditions are summarized by a state vector, k_t, which is common knowledge. Finally, in each period, each player chooses a J vector of actions or controls, $A_{bt} = [A^j_{bt}], j = 1, \ldots, J$. Let the matrix of choices be $A_t = [A_{1t}, \ldots, A_{Bt}]$.

choices, see Pinkse, Slade, and Brett (2002) and Pinkse and Slade (2004).

[4] A firm thus manufacturers one brand. However, the model could be extended to handle multi-product firms.

[5] When T is finite, we assume that the period of the data is sufficiently far from the end of the game so that players use time–independent or stationary strategies.

As is standard in the dynamic discrete–choice literature, player heterogeneity is captured by private information, u_{bt}, which in our case is a J vector, $[u^j_{bt}]$. This information, which is revealed before actions are chosen, is observed only by firm b and not by firm $b', b' \neq b$, or by the econometrician. Private information might be associated with manufacturing, retailing, or promotional cost. What is important is that, due to the presence of private information, seemingly identical players facing the same conditions can choose different actions. Furthermore, unlike previous researchers we allow private information to be correlated across players.

Correlation is potentially an important issue since, with our application, a group of B brands is sold in each store, and those stores belong to chains. In the application, we assume that each store is a 'market,' which means that the game is played within a store. However, information might be correlated across different brands sold in a given chain or across the same brand sold in different chains.

In addition to private information there is also public information, v_t, that is observed by the players but not by the econometrician. Let the single–period profit functions be $\pi_{bt} = \pi_b(A_t, k_t, u_{bt}, v_t)$. Expected payoffs in the dynamic game are then

$$\Pi_b = E \sum_{t=0}^{T} \delta^t \pi_{bt}, \tag{2}$$

where $\delta, 0 < \delta < 1$, is the discount factor and E is the expectation operator.

In this game, player b's strategy, σ_b, is a set of rules that map the state and that player's information into actions,

$$A^j_{bt} = \sigma^j_b(k_t, u^j_{bt}, v^j_t), \quad j = 1, \dots, J. \tag{3}$$

We use σ to denote the profile of strategies, $\sigma = (\sigma_1, \dots, \sigma_B)^T$.

We do not develop a formal model of equilibrium. Instead, we assume that σ is a pure–strategy equilibrium of the game but otherwise remain agnostic. Furthermore, when there are multiple equilibria, we assume that only one equilibrium is played and that firms expect that this equilibrium will continue to be played in the future. Under this assumption, we can let the data determine which equilibrium is played. This is a standard assumption in the literature.[6] However, if changes in the structure of the market are expected, this assumption might be suspect. In a stable environment, in contrast, it makes sense. Furthermore, when there are multiple markets, we assume that the same equilibrium is played in each market. The reasonableness of this assumption depends on the nature of the markets. When they are very similar to

[6] See, e.g., Bajari, Benkard, and Levin (2004).

one another, as they are in our application, it seems reasonable. However, when the structure is very different across markets, it is less plausible.

For any time period, t', the continuation payoff is defined as

$$\Pi_{bt'} = E^{\sigma}_{t'} \sum_{t=t'}^{T} \delta^{(t-t')} \pi_{bt}, \tag{4}$$

where the subscript on the expectation operator denotes the period in which the expectation is taken, and the superscript indicates that players expect the equilibrium, σ, to be played in future periods. In what follows, the superscript is suppressed.

It may help to keep the application in mind. In that context, the players are manufactures of brands of a differentiated product who choose a price, p, and whether to advertise, a. With this example, $J = 2$,[7] price is a continuous variable, $p \geq 0$, a is a zero/one variable, and the state vector consists of choices in the previous period as well as demand and cost conditions. In what follows, we assume that the advertising choice is made first. This means that each firm's pricing decision is conditional on its advertising choice. This timing assumption is not necessary, but it simplifies matters. Furthermore, as we discuss below, it seems sensible in the context of our application.

The latent–variable model is then

$$a^*_{bt} = E(\Pi_{bt}|a_{bt} = 1, \mathcal{I}_{bt}) - E(\Pi_{bt}|a_{bt} = 0, \mathcal{I}_{bt}), \tag{5}$$

and the observed–choice equation is

$$a_{bt} = \sigma^a_b(k_t, u^a_{bt}, v^a_t) = I[a^*_{bt} > 0], \tag{6}$$

where $I(.)$ is the indicator function that equals one when its argument is true and zero otherwise. In other words, players compare expected payoffs in the dynamic game when they advertise today to those when they do not, taking into account all ramifications of current choices on future profits.[8]

We would like to base our estimation on equation (6). However, we do not know the functional form of a^*_{bt}. Nevertheless, we know that it depends on expected rival

[7] With the application, $J = 3$, since we also model the use of displays.

[8] The payoff function can be written in value–function notation as

$$\Pi_{bt} = E_t\{\pi_b[\sigma(k_t, u_t, v_t), k_t, u_{bt}, v_t) + \delta V_b[k_{t+1}|k_t, \sigma(k_t, u_t, v_t)]\},$$

where V_b is the expected value of the continuation game that begins in the next period. This equation shows that players compare the expected difference in both current profits and continuation values.

choices as well as on the current state, k. We therefore approximate a^*_{bt}, either parametrically or nonparametrically.[9]

The continuous choice of p_{bt} is slightly different. Instead of comparing discrete values of expected payoffs, firms simultaneously choose prices to maximize expected payoffs, assuming that σ will be played in future periods. In addition, period–t price choices are conditional on period–t advertising choices. The first–order conditions for those choices can be manipulated to yield

$$p_{bt} = \sigma^p_b(a_t, k_t, u^p_{bt}, v^p_t) = g_b(a_t, k_t, u^p_{bt}, v^p_t). \tag{7}$$

As with the first choice, $g_b(.)$ can be approximated either parametrically or nonparametrically.

Equations (6) and 7) are our estimating equations for the firms' choice rules or policy functions. In addition, we estimate a demand equation. With that equation, sales of brand b depend on current as well as lagged choice variables.

4 The Econometric Model

There are B brands or firms, S stores or markets, and T time periods for a total of $n = BST$ observations. Firms make their advertising decisions first and subsequently make display and pricing decisions. Firm b bases its advertising decision on the value of the latent variable a^*_{bst}, which can be interpreted as the difference in expected profit between advertising and not advertising its brand at time t in store s. Firm b has private information u_{bst} and information v_{st} that is shared by all firms that sell their brands in store s in period t. As mentioned earlier, firm b does not observe rivals' advertising choices but instead optimizes against their probability of advertising.

Below we describe our estimation method. Our description does not include a theoretical justification in the form of theorems and proofs, since the asymptotic results (and the necessary conditions) follow from application of the theoretical results in Pinkse, Shen and Slade (2004). What is needed, beyond what is described below, is a mild weak–dependence condition across time and space.

The latent–variable advertising equation for brand b at time t in store s is[10]

$$a^*_{bst} = \alpha \sum_{b' \neq b} P_{b'st}(v_{st}, \theta) + x'_{bst}\beta + \gamma v_{st} - u_{bst}, \tag{8}$$

[9] When the state vector is large, only a simple approximation can be used.

[10] Notice that only the sum of probabilities matters. In particular, when those probabilities are one, only the number of rivals who choose to advertise matters, not their identities. This assumption is fairly standard.

where $\theta = [\beta', \alpha, \gamma]'$ is an unknown parameter vector, x_{bst} is a regressor vector, v_{st}, u_{bst} are standard normal errors and $P_{b'st}(v_{st}, \theta)$ is a rival probability of advertising for known v_{st}. The x_{bst}–vector includes the state, k_t, (e.g., the exogenous variables and lagged choices of all players) and can include store, brand, and time dummies. Let X_{st} be a matrix with x'_{bst} as the b–th row. As discussed earlier, v_{st} is observed by all firms, but u_{bst} is observed only by the firm producing brand b. We further assume that u_{bst} is independent across b and that u_{bst}, v_{st}, X_{st} are mutually independent. Thus there is error dependence between brands in the same store, but only through the v_{st}–terms. Note that we have not made any assumptions about potential dependence across stores or time.

Now, if v_{st} were known to the econometrician, the probability of advertising could be determined by

$$P_{bst}(v, \theta) = P(a^*_{bst} \geq 0 | v_{st}, X_{st}; \theta) = \Phi\left(\alpha \sum_{b' \neq b} P_{b'st}(v_{st}, \theta) + x'_{bst}\beta + \gamma v_{st}\right). \quad (9)$$

For each s, t, there are B equations of the form (9), which can be jointly (and numerically) solved to determine the $P_{bst}(v, \theta)$. Uniqueness of the solution follows immediately from Banach's fixed point theorem, provided that the value of α is not too large. It can be shown that a sufficient condition for identification is that $|\alpha| < 1/\big(\phi(0)(B-1)\big)$, where $\phi(\cdot)$ is the standard normal density function. This condition can be reasonable when B is small (like in our case), but for large B this condition can be restrictive.

When $B = 2$, joint choice probabilities are given by

$$P(a_{bst} = 1, a_{b'st} = 1 | X_{st}) = P^{11}_{bb'st}(\theta) = \int P_{bst}(v, \theta) P_{b'st}(v, \theta)\phi(v)dv, \quad (10)$$

where the remaining three combinations are defined analogously.[11] We use the loglikelihood function under independence across time and stores and correct for dependence ex post. This is less efficient than full ML if the complete dependence structure were known, but unlike full ML it requires no precise knowledge of the dependence structure. The loglikelihood function is

$$\hat{L}(\theta) = \sum_{bst} \ell_{bb'st}(\theta), \quad (11)$$

where

$$\begin{aligned}\ell_{bb'st}(\theta) = {} & a_{bst}a_{b'st} \log P^{11}_{bb'st}(\theta) + a_{bst}(1 - a_{b'st}) \log P^{10}_{bb'st}(\theta) \\ & + (1 - a_{bst})a_{b'st} \log P^{01}_{bb'st}(\theta) + (1 - a_{bst})(1 - a_{b'st}) \log P^{00}_{bb'st}(\theta). \quad (12)\end{aligned}$$

[11] The extension for larger B is also similar.

Note that the loglikelihood will take the same value when $\gamma = \gamma^*$ and when $\gamma = -\gamma^*$, and γ must hence be constrained to be nonnegative.

The results in Pinkse, Shen and Slade (2004) now imply that

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N\Big(0, \Gamma^{-1}(\theta)\Omega(\theta)\Gamma^{-1}(\theta)\Big), \tag{13}$$

where

$$\Gamma(\theta) = \lim_{n\to\infty} n^{-1} \sum_{bb'st} E\left(\frac{\partial^2 \ell_{bb'st}}{\partial\theta\partial\theta'}(\theta)\right), \quad \Omega(\theta) = \lim_{n\to\infty} n^{-1} \sum_{bb'st\tilde{b}\tilde{b}'\tilde{s}\tilde{t}} E\left(\frac{\partial \ell_{bb'st}}{\partial\theta}(\theta)\frac{\partial \ell_{\tilde{b}\tilde{b}'\tilde{s}\tilde{t}}}{\partial\theta'}(\theta)\right), \tag{14}$$

provided that the limits exist. Loosely speaking, these limits typically exist if the sum over all covariances increases at a rate of n.

Ω can be estimated using the procedure proposed in Pinkse, Shen and Slade (2004). An estimator for Γ is

$$\hat{\Gamma}(\hat{\theta}) = n^{-1} \sum_{bb'st} \frac{\partial^2 \ell_{bb'st}}{\partial\theta\partial\theta'}(\hat{\theta}). \tag{15}$$

To compute standard errors one hence needs expressions for the first and second partials of ℓ and therefore also of P. From (10) it follows that

$$\frac{\partial P^{11}_{b'bst}}{\partial\theta}(\theta) = \int \left(\frac{\partial P_{bst}}{\partial\theta}(v,\theta)P_{b'st}(v,\theta) + P_{bst}(v,\theta)\frac{\partial P_{b'st}}{\partial\theta}(v,\theta)\right)\phi(v)dv, \tag{16}$$

$$\begin{aligned}\frac{\partial^2 P^{11}_{bb'st}}{\partial\theta\partial\theta'}(\theta) = \int \Bigg(&\frac{\partial^2 P_{bst}}{\partial\theta\partial\theta'}(v,\theta)P_{b'st}(v,\theta) + \frac{\partial^2 P_{b'st}}{\partial\theta\partial\theta'}(v,\theta)P_{bst}(v,\theta) \\ &+\frac{\partial P_{bst}}{\partial\theta}(v,\theta)\frac{\partial P_{b'st}}{\partial\theta'}(v,\theta) + \frac{\partial P_{b'st}}{\partial\theta}(v,\theta)\frac{\partial P_{bst}}{\partial\theta'}(v,\theta)\Bigg)\phi(v)dv. \end{aligned} \tag{17}$$

In an appendix we develop a procedure to recover the first and second partials of P_{bst}.

5 The Data and Market

The game that we model is played by firms that produce brands of a differentiated product that is sold in retail outlets. In each period, each decision maker must choose a price and must make promotional decisions for a single brand. Since the decisions that are made today affect future own and rival decisions and sales, the game is dynamic.

5.1 Data Sources

The data that we use were collected by Information Resources, a marketing firm in Chicago. These data pertain to retail sales of grocery-store products in 1984–1985. Information Resources selected several small US towns for data–collection purposes. Each town was required to have no more than ten grocery stores and to have no neighboring town within 100 miles.

Approximately ten percent of the households in each town were randomly contacted and asked to participate in the experiment. Agreeing households were given a card to use when shopping. With the aid of this card, all of the household's grocery–store purchases were recorded electronically. The final sample for each town consists of approximately five percent of all households — those who used their cards at least once every eight weeks during the two-year data-collection period.

The household data contain a complete set of demographic variables, each household's purchase record for the two-year period, and information on coupons redeemed. For the purpose of this study, however, household demographics are not particularly relevant. For this reason, we converted the household–purchase data into sales data by week, store, and brand of product. Moreover, we recorded the number of manufacturer coupons redeemed by brand, store, and week.

In addition to household demographics and purchases, Information Resources compiled weekly data by product (disaggregated by brand) and by grocery chain (disaggregated by store). These data consist of price, information about advertising in local newspapers, and whether the product was featured in an aisle display.

The subset of the data that is used in the current study pertains to sales of saltine crackers in Williamsport PA.[12] Saltines have a number of advantages. Perhaps the most important is that a saltine cracker is a relatively homogeneous product that is sold almost exclusively in two-pound boxes of a standard shape. In addition, there is very little national advertising for saltines and almost all promotion is local.

There are three national brands of saltines: American Brands – Sunshine, Keebler – Zesta, and R.J.R. Nabisco – Premium, as well as many generic or private–label brands. In constructing the data, the private–labels were combined into one generic brand. For each brand b, store s, and week t, the constructed variables are.

Sales (Q) purchases by households (number of two–pound boxes),

Advertising dummy (AD) = 1 if the brand is advertised this week,

Rival advertising (RAD) the number of rival brands that are advertised,

Time elapsed (TLAD) in weeks since the brand was last advertised,

Price (PRICE) in dollars per box,

[12] These data are also used in Slade (1995 and 1998).

Rival price (RP) a share–weighted average across rivals,

Display dummy (DISP) = 1 if the brand was on display,

Rival displays (RDISP) the number of rival brands that were on display,

Manufacturer coupons (COUP) the number of coupons redeemed,

for $b = 1, ..., B = 4$, $s = 1, ..., S = 10$, and $t = 1, ..., T = 104$. The number of observations in each cross section is thus 40 whereas the number of time periods, T, is 104, which leads to 4160 observations.

Cost variables are more aggregate and sampled at less frequent intervals. Two production-factor prices were collected. Those prices are average hourly earnings in cracker and cookie manufacturing (WAGE) and the producer-price index for intermediate foods and feeds (PIF). Both are published by the U.S. Bureau of Labor Statistics (BLS). The BLS also publishes a monthly wholesale price for saltine crackers (PWHOLE), an average across brands. This is a list price that does not reflect the discounts that the retailers receive. Nevertheless, we use this variable to reflect cost changes that are not captured by the factor prices. One advertising–factor price was collected. It is the Bates Media Survey cost index per 1000 viewers, newspapers (CNAD), which is published in Broadcasting magazine.

On the demand side, an income variable (EARN) was constructed as the product of nonagricultural employment and average weekly earnings per employed person. Both of those series, which also come from the BLS, are for Williamsport PA.

To convert the monthly cost and income variables into weekly data, the following procedure was adopted. First each variable in week t, x_t, was given the value of that variable in the corresponding month. Each series was then smoothed using the linear filter, $x'_t = 0.25x_{t-1} + 0.50x_t + 0.25x_{t+1}$. With this filter, the value of the weekly variable in the first (last) week of each month is a convex combination of the current and previous (following) month's values.

Finally, all monetary variables (product and factor prices and income) were converted into real variables by dividing by the US consumer–price index, all items (CPI).

5.2 Preliminary Data Analysis

Table 1 contains summary statistics for the brands. The first two rows show averages for the entire sample and for the three major brands, respectively, whereas the remaining rows contain statistics that are disaggregated by brand. The table shows that the private-label group of brands, which adopts a low-price policy, jointly sells the largest number of boxes. In contrast, the largest single brand, Nabisco Premium, is the highest priced. On average, a given brand is advertised in a local newspaper 6% and displayed in a given store 12% of the time. However, there are a number of differences across brands. In particular, the private label brands advertise relatively infrequently but use displays much more often than the major brands. Finally, the vast majority of manufacturer coupons that are redeemed are for one brand — Nabisco Premium.

It is clear that the private–label brands use different pricing and promotional policies from the majors. In particular, they rely on low prices rather than heavy advertising to promote their products. In addition, since sales of individual private labels are relatively small, the manufacturers of those products are less apt to behave in a sophisticated strategic fashion. For this reason, in our empirical analysis, we limit attention to the strategies of the major–brand manufacturers. Our analysis, however, is conditioned on the prices of the private labels.

Williamsport has ten grocery stores that belong to four chains: two regional, one national, and one independent. Pricing decisions are often (but not always) uniform within a chain. Table 2, which contains summary statistics by chain, shows that a regional retailer accounts for approximately one half of the market and that, even though the numbers of stores owned by each of the remaining chains are different, their market shares are approximately equal. Pricing is more uniform across chains than across brands.[13] Nevertheless, the largest chain sets the highest prices. Moreover, there are substantial differences in the chains' use of promotions. In particular, chain three advertises and uses displays much less frequently than the others.

Since the model is dynamic, the first three weeks in the sample were dropped so that initial values of the variables could be constructed. We therefore have 30 observations per cross section and 101 weeks or 3030 observations that are used in the empirical analysis.

Table 3 contains correlation coefficients for the choice variables, advertising (AD), price (P), and displays (DISP), as well as for the number of coupons redeemed (COUP). The table shows that newspaper ads, low prices, and display activity tend to occur together. However, the number of coupons that are redeemed is not correlated with the strategic variables. The latter is true because, even though the manufacturer issues the coupons, the consumer decides when to redeem them.

5.3 Product Rivalry

Each week, the manufacturers' local marketing managers offer chain managers 'deals'. These consist of price recommendations and advertising and/or display subsidies. Chain managers then choose a 'package', a balanced menu of items to feature. *A priori*, it is not clear where the competition lies and who makes the important decisions.

[13] The prices in table 2 are averages across the three major brands.

Prior to the statistical analysis, grocery-chain marketing managers were interviewed and questioned about their policies. A fairly coherent story emerged from the process. The managers alleged that less than ten percent of households contain comparison shoppers who visit several stores in a week to search for the lowest-priced items. The remaining 90 percent frequent the same store in most weeks. Their choice of store is determined by location (often proximity to work or home) and by the quality of the store (freshness of produce and meat, product offerings, overall pricing policies, etc.). The anecdotal evidence therefore favors a model where competition is among brands within a store. This does not mean, however, that chains fail to compete. Rather it implies that chains compete via their total offerings rather than through individual items such as saltine crackers.

We therefore model each store as a 'market,' and our previous assumptions imply that, if the game has multiple equilibria, the same one is played in each store. Given the similarity of our 'markets,' this assumption seems reasonable.

We assume that the decision makers are the three manufacturers of major brands of saltines and go on to examine the strategies that they use. In other words, we model manufacturers as price and promotion setters whereas retailers are assumed to be more passive. Clearly, the true situation is more complex than the one modeled. Nevertheless, this simplification appears to be more realistic than the opposite extreme of passive manufacturers. Manufacturers are assumed to anticipate how their choices of prices and promotional activities will affect expected current and future brand sales and revenues.

The prices that we use are those that the retailer, not the manufacturer, receives. We assume that the manufacturer maximizes the joint surplus, the discounted flow of the brand's retail revenue minus manufacturing and retailing costs. Since transactions between manufacturers and retailers are not completely arm's length, the surplus can be distributed between the two in an agreed–upon fashion. For example, the wholesale price plus fixed fees, such as positive or negative slotting allowances, can be used for this purpose.

We make use of a number of symmetry assumptions in modeling the manufacturers' strategies. In particular, we assume that only the number of rivals who advertise or display their brands in a given store and week matters, not the identities of those manufacturers.[14] In addition, we use a share-weighted average rival price, where the shares are those shown in table 1. The brands and chains, however, are clearly not symmetric. For example, in spite of the fact that Nabisco Premium charges the highest prices, it has the largest share of the market. Furthermore, tables 1 and 2 indicate that there are many other differences in promotion and pricing policies across brands

[14] This is a common assumption in the literature (e.g., Aguirregabiria and Mira, 2004).

and chains. Such time–invariant unobserved heterogeneity is captured by brand and chain fixed effects. Time–varying unobserved heterogeneity, in contrast, is captured by the decision makers' private information.

Finally, advertising occurs less frequently than price changes or displays occur. In addition, preliminary data analysis revealed that advertising has a large impact on sales, much larger than, for example, featuring a product in an aisle display. Furthermore, the cost of advertising is higher than the cost of changing price or activating a display. For these reasons, we feel comfortable with the specification outlined earlier in which advertising is the primary decision that is made first.

6 Empirical Results

6.1 Preliminary Estimations

Table 3 contains results from some preliminary estimations. The choice rules are based on equations (6) and (7).[15] We assume that equation (7) explains the manufacturers' price recommendations, and that there can be differences between those recommendations and the retailers' decisions. We append an additive error to capture those differences.

We also present a demand equation so that the dynamic effects of the firms' choices can be assessed. The four equations were estimated by ordinary probit (the equations for advertising and displays), OLS (the price equation), and two–stage least squares (the demand equation).[16] In the table, a subscript minus one is used to denote the value of a variable in the previous week.

The first equation, which is a linear approximation to equation (6), is a preliminary version of the advertising strategy.[17] The dependent variable, AD, is the advertising dummy. In addition to the brand and chain fixed effects, this equation contains only the current state variables — current values of the exogenous demand and cost variables and lagged choice variables. We assume that players form expectations concerning rival choices using the information that they currently have (i.e., k_t) but we do not model expectation formation. The table shows that a manufacturer is less apt to advertise if a rival advertised his brand in the last period. In addition, the longer the time that has elapsed since a brand was last advertised, the more likely it is to be advertised today. Finally, advertising is less likely to occur when prices were

[15] We assume that the display choice is also conditioned on the advertising decision and therefore that equation contains the current state variable AD. Otherwise, the equation is identical to (6).

[16] We use as instruments the exogenous variables in the system of equations as well as appropriately lagged values of endogenous variables.

[17] We use a linear approximation because the state vector is large.

low in the previous period.

The pricing and display equations contain the same variables as the advertising equation with one addition. Since we assume that the advertising decision is made first, those equations also contain the variable AD that indicates whether or not the brand will be advertised in this period.

First, consider the pricing equation. It is clear that manufacturers choose a lower price in periods when they plan to advertise. However, prices are higher immediately after a brand has been advertised or featured on display. In addition, prices are lower after rivals have advertised or charged high prices. Finally, prices are positively autocorrelated and increase as the time since the brand was last advertised lengthens.

Now consider the display equation. Brands are also more likely to be featured on display if they will be advertised this period. In addition, displays are more likely in periods immediately after any brand was advertised. Finally, display activity is positively autocorrelated and becomes more likely as the time since the brand was last advertised lengthens.

A picture emerges in which advertising, display activity, and price reductions are complementary activities. However, brands tend to be advertised for only one week, whereas low prices and displays tend to persist for a while after the ad has been discontinued.

The final equation in table 3 is the demand equation. The endogenous variables in that equation are the current values of own and rival choice variables (AD, P, and DISP) as well as the number of coupons redeemed.[18] In addition to the exogenous variables, we use appropriately lagged values of the endogenous variables as instruments. The table shows that sales increase when advertising occurs, when brands are featured, when prices are low, and when coupons are issued. Sales are lower, however, in periods following own or rival advertising campaigns. This result could be due to an inventory effect. Indeed, households that have just purchased a box of crackers are less apt to do so in the immediate future. Finally, sales fall as the time since a brand was last advertised lengthens.

6.2 Spatial Estimations

We computed estimates of the advertising decision rule as indicated in section 4.[19] The results can be found in table 5, which compares ordinary and spatial probits. Two versions of each type are shown. The first sets α (the coefficient of current rival

[18] COUP is an endogenous variable because consumers can only redeem a coupon when they purchase the product. There is thus feedback from sales to coupons.

[19] Since the objective function is not globally concave, we tried a number of different starting values.

advertising) equal to zero *a priori*, whereas the second allows advertising choices to be simultaneous. Furthermore, the two probits with $\alpha = 0$ differ in that the standard errors of the coefficients of the spatial probit were calculated using the technique that is developed in Pinkse, Shen, and Slade (2004). In other words, the errors are corrected for spatial and time–series correlation of an unknown form.

Both spatial probits in table 5 were estimated under the assumption that γ equals zero. We also estimated two models in which γ was allowed to vary, one assuming $\alpha = 0$, the other allowing α to vary as well, but the estimates of γ were identically zero, presumably because at no time were two brands advertised simultaneously in the same store.[20]

Comparing the probits in columns 1 and 3 of table 5 (i.e., those with $\alpha = 0$), one can see that the spatial–probit t–statistics are larger in magnitude. This difference is due to the negative correlation in advertising decisions across time and brands, over and above what is explained by the model.

Comparing the two spatial probits (columns 3 and 4), the table shows that the coefficients are often larger in magnitude when α is not zero. The standard errors, in contrast, are sometimes larger and sometimes smaller. Nevertheless, the two equations are quite similar to one another.

The comparison between the ordinary and spatial probits with $\alpha \neq 0$ (columns 2 and 4) is the most interesting. Indeed, although all of the coefficients in the two equations have the same sign and most are similar in magnitude, the estimates of the parameter α itself are very different across equations. In particular, when the simultaneity problem is dealt with in a rigorous fashion, the magnitude of the estimate of α drops substantially, which is evidence of a negative bias.[21] A negative α suggests that, if a firm thinks that the probability that a rival will advertise is high, for strategic reasons, that firm is less likely to advertise itself. A negative bias, however, suggests that this strategic effect is exaggerated when the simultaneity problem is ignored.

It is not possible to determine the source of the negative bias in the estimate of α definitively. However, a negative bias is consistent with the following story, which we think is plausible.

When searching for the source of a bias, one normally considers the possibility that excluded variables are correlated with the included variables. In particular, we would like to find an omitted variable that is correlated with rival advertising

[20] The t–statistics in the full model would typically be smaller because of the additional unknown coefficient. How much smaller is difficult to establish since zero is at the boundary of the parameter space, a case not covered by standard asymptotic analysis, but see e.g. Hirano and Porter (2004).

[21] We phrase our conclusion cautiously because the t statistic of the estimate of α in the spatial probit corresponds to a p–value of 17% (two–sided), which is above any commonly used significance level.

decisions. The most obvious omitted variable is the level of each brand's inventories in the store or chain, and failure to include that variable could cause a negative bias.[22] To illustrate, consider an example with two brands. If the size of the market is relatively constant and advertising merely shifts sales from one brand to the other, when inventories of one brand are high, inventories of the other will tend to be low. This means that, when it is advantageous for one firm to advertise in order to reduce its inventories, the other firm is apt to find that advertising is disadvantageous, at least for inventory–reduction motives. This means that brand–level inventories are a common causal factor that can cause a bias in the estimate of α. Furthermore, that bias will be negative.

We experimented with a number of alternative specifications of the advertising equation. In particular, we constructed the rival advertising variable with and without the private labels, and we estimated versions of the equation with and without the coupon variable. However, in all cases, the ordinary probit estimate of the coefficient of current rival advertising was substantially larger than the spatial estimate shown in table 5. We conclude that the size difference is robust.

6.3 Comparative Dynamics

We can use the estimated model for at least two purposes. First, we can evaluate its ability to reproduce the dynamics that characterize the data. And second, if the first exercise is satisfactory, we can perform experiments. In particular, we can evaluate the long–run implications of advertising competition for prices and sales.

For simplicity, we performed the simulations with only two symmetric firms.[23] Furthermore, we set all exogenous variables to their means and added their effects to the constant terms.

The basic version of the model reproduces the advertising dynamics well. Pricing dynamics, in contrast, are captured less well, but modeling prices was not our objective. With the base case, firms advertise infrequently, and their advertising efforts are staggered.[24] Furthermore, an advertising campaign is accompanied by a low price. However, price jumps up just after a campaign and increases gradually until the next campaign. Finally, sales increase dramatically when a product is advertised but fall just as dramatically immediately afterwards, after which they begin to ascend gradually.

[22] Unfortunately, we have no data on store or chain inventories.

[23] This exercise is simplified by the fact that, in the data, two rivals never advertise at the same time. The rival–advertising variable is therefore dichotomous. In addition, we replaced the average–rival price with a single–rival price.

[24] Since our simulations are nonstochastic, eventually the model settles into regular cycles.

Our experiments involve varying the frequency of advertising. To do this, we vary the threshold that triggers campaigns. With the base case, we assume that advertising occurs whenever the probability of advertising is greater than 0.5. With other cases, we allow this number to change.

It is difficult to predict *a priori* how advertising will affect prices and sales in the dynamic game. For example, consider the impact on prices. The direct effect on own price is negative. Indeed, price is lower in the period when the product is advertised and rises in the next period but by a lesser amount. However, indirectly, rival price falls next period, which causes own price to rise, which in turn causes rival price to fall, and so forth. Finally, advertising is associated with display activity, and prices rise after a product has been on display.

We find that when only one firm advertises more frequently, on average its price is lower and its sales are higher. When both advertise more frequently, however, sales are higher as before, but prices are no longer lower. This reversal occurs because rival advertising is associated with lower rival prices, which cause own prices to rise. Advertising is thus expansionary, but consumers pay higher prices.

7 Conclusions

We have developed a dynamic discrete–choice estimator that can be used in strategic contexts. It is particularly useful in situations where the state vector is large and when decisions and/or private information can be correlated. The state vector can be large because, for example, there are multiple firms, each of which chooses multiple instruments of rivalry.[25] Decisions can be correlated because firms behave in a strategic manner. Finally, information can be correlated because, for example, many different brands are sold by the same chain and the same brand is sold in many chains. In other words, correlated private information can arise on both the manufacturing and the retailing side of the market.

We apply our estimator to evaluate the long–run effects of promotional decisions in an industry that produces an established standardized product. In other words, there is little technical change or market turbulence in our industry — the saltine–cracker industry. We find that although firms most likely benefit from advertising, the effect on consumers is more ambiguous.[26]

We conclude that when firms advertise, they do not simply steal the customers

[25] It can also be large because firms produce multiple products, a situation that we do not consider in our application.

[26] We do not perform a formal welfare analysis for two reasons. First, we are working in a very partial–equilibrium setting, and second, our model is not structural.

of their rivals. Instead, the market as a whole expands. This expansionary effect is perhaps due to the informative nature of grocery–store promotions.[27] However, consumers must pay for their better information in the form of higher prices.

There are other questions, such as the effect on consumer welfare of the introduction of new brands, that we cannot address in the context of our model. Indeed, that exercise requires a solution to a structural dynamic discrete–choice game, which is precisely what our estimator avoids. Its virtue is that it allows us to estimate players' decision rules in a relatively straightforward way. In particular, we do not have to specify functional forms and distributions for all of the model primitives or to optimize over the full vector of parameters using, for example, simulated maximum likelihood techniques. The computational burden of such estimations is prohibitive when the state vector is large and information is correlated.

[27] Clearly, the expansion cannot go on forever. At some point, consumers will become saturated with information, which could even have a negative impact on sales.

References Cited

Ackerberg, D. (2003) "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination," *International Economic Review*, 44: 1007–1040.

Aguirregabiria, V. and Mira, P. (2004) " Sequential Estimation of Dynamic Discrete Games," Boston University mimeo.

Anand, B. and Shachar, R. (2001) "Advertising, the Matchmaker," Harvard Business School mimeo.

Bajari, P., Benkard, C.L., and Levin, J. (2004) "Estimating Dynamic Models of Imperfect Competition," Stanford University mimeo.

Chamberlain, G. (1992) "Efficiency Bounds for Semiparametric Regression," *Econometrica* 60: 567–596.

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

Goeree, M.S. (2004) "Advertising in the US Personal Computer Industry," Claremont McKenna College mimeo.

Hirano, K. and J.R. Porter (2003) "Asymptotic Efficiency in Parametric Structural Models with Parameter-Dependent Support," *Econometrica* 71, 1307–1338.

Lucas, R. (1976) "Econometric Policy Evaluation: A Critique," in *The Phillips Curve and Labor Markets*, K. Brunner and A, Meltzer (eds.) Amsterdam: North Holland.

Manski, C.F. (1975) "Maximum Score Estimation of the Stochastic Utility Model of Choice," *Journal of Econometrics* 3: 205–228.

Pakes, A., Ostrovsky, M., and Berry, S. (2004) "Simple Estimators for the Parameters of Discrete Dynamic Games," Harvard University mimeo.

Pesendorfer, M. and Schmidt–Dengler, P. (2003) "Identification and Estimation of Dynamic Games," London School of Economics mimeo.

Pinkse, J., Slade, M.E., and Brett, C. (2002) "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70: 1111–1155.

Pinkse, J. and Slade, M.E. (2004) "Mergers, Brand Competition, and the Price of a Pint," *European Economic Review*, 48: 617–643.

Pinkse, J., Shen, L., and Slade, M.E. (2004) "Dynamic Spatial Probit with Fixed Effects Using One–Step GMM: An Application to Mine Operating Decisions," University of Warwick mimeo.

Seim, K. (2004) "An Empirical Model of Firm Entry with Endogenous Product–Type Choices," Stanford University mimeo.

Shum, M. (2004) "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast–Cereals Market," *Journal of Economics and Management Strategy*, 13: 241–272.

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

Slade, M.E. (1998) "Optimal Pricing with Costly Adjustment: Evidence from Retail–Grocery Prices," *Review of Economic Studies*, 65: 87–107.

Slade, M.E. (2004) "The Role of Economic Space in Decision Making," *Annales d'Economie et de Statistique*, forthcoming.

Tamer, E. (2003) "Incomplete Simultaneous Discrete Response Models with Multiple Equilibria," *Review of Economic Studies*, 70: 147–165.

A Partial derivatives

We drop the dependence on s, t in the notation and define $\Delta_b(v,\theta) = \alpha \sum_{b' \neq b} P_{b'}(v,\theta) + x_b'\beta + \gamma v$, such that

$$P_b(v,\theta) = \Phi\big(\Delta_b(v,\theta)\big), \tag{18}$$

$$\frac{\partial P_b}{\partial \theta}(v,\theta) = \frac{\partial \Delta_b}{\partial \theta}(v,\theta)\phi\big(\Delta_b(v,\theta)\big), \tag{19}$$

$$\frac{\partial^2 P_b}{\partial \theta \partial \theta'}(v,\theta) = \left(\frac{\partial^2 \Delta_b}{\partial \theta \partial \theta'}(v,\theta) - \Delta_b(v,\theta)\frac{\partial \Delta_b}{\partial \theta}(v,\theta)\frac{\partial \Delta_b}{\partial \theta'}(v,\theta)\right)\phi\big(\Delta_b(v,\theta)\big). \tag{20}$$

$$\frac{\partial \Delta_b}{\partial \theta}(v,\theta) = \mu_b(v,\theta) + \alpha \sum_{b' \neq b} \frac{\partial P_{b'}}{\partial \theta}(v,\theta), \tag{21}$$

with

$$\mu_b(v,\theta) = \begin{bmatrix} \sum_{b' \neq b} P_{b'}(v,\theta) \\ v \\ x_b \end{bmatrix},$$

rearranging terms yields

$$\frac{\partial P_b}{\partial \theta}(v,\theta) - \alpha\phi_b(v,\theta) \sum_{b' \neq b} \frac{\partial P_b'}{\partial \theta}(v,\theta) = \mu_b(v,\theta)\phi_b(v,\theta), \quad b = 1, \dots, B, \tag{22}$$

which can be solved for the $\partial P_b / \partial \theta$–quantities by writing

$$\begin{bmatrix} I & -\alpha\phi_1 I & \cdots & -\alpha\phi_1 I \\ -\alpha\phi_2 I & I & \ddots & -\alpha\phi_2 I \\ \vdots & \ddots & I & \vdots \\ -\alpha\phi_B I & \cdots & -\alpha\phi_B I & I \end{bmatrix} \begin{bmatrix} \frac{\partial P_1}{\partial \theta} \\ \vdots \\ \frac{\partial P_B}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \mu_1\phi_1 \\ \vdots \\ \mu_B\phi_B \end{bmatrix}. \tag{23}$$

Similarly,

$$\frac{\partial^2 \Delta_b}{\partial \theta \partial \theta'}(v,\theta) = \Upsilon_b(v,\theta) + \alpha \sum_{b' \neq b} \frac{\partial^2 P_{b'}}{\partial \theta \partial \theta'}(v,\theta), \tag{24}$$

with (omitting arguments)

$$\Upsilon_b = \sum_{b' \neq b} \Big(\frac{\partial P_{b'}}{\partial \theta}\iota' + \iota\frac{\partial P_{b'}}{\partial \theta'}\Big),$$

where $\iota = [1, 0, \dots, 0]'$. Thus, if $\Upsilon_b^* = \Upsilon_b - \Delta_b(\partial \Delta_b/\partial \theta)(\partial \Delta_b/\partial \theta')$, then

$$\frac{\partial^2 P_b}{\partial \theta \partial \theta'} = \left(\Upsilon_b^* + \alpha \sum_{b' \neq b} \frac{\partial^2 P_{b'}}{\partial \theta \partial \theta'}\right)\phi_b. \tag{25}$$

Thus, the second partials can be solved for by using

$$\begin{bmatrix} I & -\alpha\phi_1 I & \cdots & -\alpha\phi_1 I \\ -\alpha\phi_2 I & I & \ddots & -\alpha\phi_2 I \\ \vdots & \ddots & I & \vdots \\ -\alpha\phi_B I & \cdots & -\alpha\phi_B I & I \end{bmatrix} \begin{bmatrix} \frac{\partial^2 P_1}{\partial\theta\partial\theta'} \\ \vdots \\ \frac{\partial^2 P_B}{\partial\theta\partial\theta'} \end{bmatrix} = \begin{bmatrix} \Upsilon_1^* \phi_1 \\ \vdots \\ \Upsilon_B^* \phi_B \end{bmatrix}. \qquad (26)$$

Table 1: **Summary Statistics by Brand**

Brand	Mkt Share	Price (P)	Advertising (AD)	Display (DISP)	Coupons (COUP)
	%	$	%	%	#
All Brands	100	1.03	5.9	12.1	7.1
3 Major	55	1.15	6.2	9.3	9.5
American Brands (Sunshine)	9	1.11	8.4	10.8	0.2
Keebler (Zesta)	8	1.17	3.5	5.7	0.2
Nabisco (Premium)	38	1.18	6.7	11.3	28.0
Private Label	45	0.68	4.9	20.5	0.1

Table 2: **Summary Statistics by Chain**

Chain	Stores	Type	Mkt Share	Price (P)	Advertising (AD)	Display (DISP)	Coupons (COUP)
			%	$	%	%	#
1	4	Regional	54	1.19	7.9	9.1	9.2
2	1	Regional	15	1.11	3.7	17.8	22.2
3	2	National	14	1.11	1.0	5.4	10.7
4	3	Independent	17	1.15	8.3	9.2	4.7

Table 3: **Correlation Coefficients**

	AD	P	DISP	COUP
AD	1.00	-0.44**	0.44**	0.09
P	—	1.00	-0.33**	-0.08
DISP	—	—	1.00	0.06
COUP	—	—	—	1.00

** denotes significance at 1%.

Table 4: **Preliminary Estimations**

Dep. Var.	Advertising (AD)	Price (P)	Display (DISP)	Demand (Q)
AD		-0.257	2.034	201.9
		(-39.1)	(17.9)	(4.3)
AD_{-1}	-0.189	0.087	0.357	-78.4
	(-1.0)	(10.8)	(2.3)	(-3.6)
RAD				-12.0
				(-1.8)
RAD_{-1}	-0.542	-0.015	0.223	
	(-3.7)	(-2.8)	(1.6)	
$TLAD$	0.031	0.0004	0.007	-0.511
	(6.8)	(3.3)	(2.5)	(-3.7)
P				-146.0
				(-3.6)
P_{-1}	-1.612	0.521	0.135	
	(-4.1)	(33.2)	(0.4)	
PR				25.6
				(0.4)
PR_{-1}	-1.430	-0.118	0.297	
	(-1.3)	(-2.7)	(0.3)	
$DISP$				75.4
				(3.2)
$DISP_{-1}$	-0.015	0.022	1.358	
	(-0.1)	(3.7)	(12.5)	
$RDISP$				-28.0
				(-6.0)
$RDISP_{-1}$	0.018	0.005	-0.081	
	(0.3)	(1.5)	(-1.1)	
$COUP$				39.3
				(21.0)
Cost Var.	yes	yes	yes	no
R^2	0.14^a	0.64	0.41^a	0.40
Technique	Probit	OLS	Probit	2SLS

3030 observations.
All equations include demand variables, a trend, and brand and chain fixed effects.
t statistics in parentheses.
[a] Cragg–Uhler R^2.

Table 5: **Advertising Equations: Ordinary and Spatial Probits Compared**

Variable	Ordinary $\alpha = 0$	Ordinary $\alpha \neq 0$	Spatial[a] $\alpha = 0$	Spatial[a] $\alpha \neq 0$
AD_{-1}	-0.189	-0.361	-0.189	-0.218
	(-1.0)	(-1.9)	(-1.2)	(-1.5)
RAD (α)		-2.474		-1.320
		(-3.3)		(-1.4)
RAD_{-1}	-0.542	-0.616	-0.542	-0.623
	(-3.7)	(-4.0)	(-4.1)	(-6.0)
TLAD	0.031	0.031	0.031	0.030
	(6.8)	(6.7)	(8.4)	(11.4)
P_{-1}	-1.612	-1.791	-1.612	-1.585
	(-4.1)	(-4.4)	(-12.2)	(-120)
PR_{-1}	-1.430	-1.672	-1.430	-2.841
	(-1.3)	(-1.4)	(-7.0)	(-4.0)
$DISP_{-1}$	-0.015	0.037	-0.015	-0.074
	(-0.1)	(0.2)	(-0.2)	(-1.1)
$RDISP_{-1}$	0.018	0.024	0.018	0.015
	(0.3)	(0.3)	(0.9)	(0.5)
Cragg–Uhler R^2	0.14	0.20		
Technique	Probit	Probit	PSS Probit	New

[a] Pinkse–Shen–Slade (2004) (PSS) t statistics in parentheses.
3030 observations.
Demand and cost variables, a trend, and brand and chain fixed effects are included.
Ordinary probit with $\alpha \neq 0$ uses realized rival advertising variables.
Spatial probit with $\alpha \neq 0$ uses expected rival advertising variables.