Structural Applications of the Discrete Choice Model[1]

JEAN-PIERRE DUBÉ (CHAIR), PRADEEP CHINTAGUNTA (CHAIR), AND AMIL PETRIN
Graduate School of Business, University of Chicago

BART BRONNENBERG
Anderson School, University of California Los Angeles

RON GOETTLER
School of Industrial Administration, Carnegie Mellon University

P. B. [SEETHU] SEETHARAMAN
Olin School, Washington University at St. Louis

K. SUDHIR
SOM, Yale University

RAPHAEL THOMADSEN
Graduate School of Business, Columbia University

YING ZHAO
Hong Kong University of Science and Technology

October 2001

Abstract

A growing body of empirical literature uses structurally-derived economic models to study the nature of competition and to measure explicitly the economic impact of strategic policies. While several approaches have been proposed, the discrete choice demand system has experienced wide usage. The heterogeneous, or "mixed", logit in particular has been widely applied due to its parsimonious structure and its ability to capture flexibly substitution patterns for a large number of differentiated products.

We outline the derivation of the heterogeneous logit demand system. We then present a number of applications of such models to various data sources. Finally, we conclude with a discussion of directions for future research in this area.

Key words: structural modeling, firm conduct, policy simulation, consumer welfare

---

[1]This paper is based on the session "Structural Models and Discrete Choice" at the 2001 Choice Symposium, hosted by UC Berkeley.

1. Introduction

Recent quantitative marketing research has turned to structural econometric models to study firm competition for differentiated products. Most of this literature has focused on describing the nature of competition and, in particular, identifying the specific "game" being played between competing firms (see Kadiyali et al. 2001 for a thorough survey). Since most of the emphasis has been on studying the supply-side, the empirical treatment of demand has typically been approximately structural, using linear and log-linear approximations to the aggregate demand function. Several recent papers have begun focusing on more rigorous specifications for the aggregate demand system that derive formally from the theory of utility-maximizing consumers facing a budget constraint. While numerous approaches for modeling aggregate demand systems for differentiated products have been developed (e.g. residual demand—Baker and Bresnahan 1985, multi-level demand system—Hausman 1997) researchers are turning increasingly to the discrete choice model. We discuss the empirical implementation of the mixed multinomial logit, or random coefficients multinomial logit, discrete choice demand system. We then present several recent applications of the model to highlight the potential benefits to marketing.

Small and Rosen (1981) demonstrate that the aggregation of standard choice models (e.g. logit and probit) generates viable economic demand systems. Typically, marketing researchers estimate the choice model using microeconomic data containing individual choices, such as the scanner panels (Guadagni and Little 1983). The demand system derives from aggregating the estimated choice model (e.g. Trajtenberg 1989). However, in instances for which only aggregate data are available, one may still be able to identify the parameters of the choice model by estimating the demand system directly (Allenby 1989; Allenby and Rossi 1991, and Berry 1994). Recent research has focused on the identification of heterogeneity in tastes when aggregate data are available (Berry, Levinson and Pakes 1995—hereafter referred to as BLP).

The use of the logit demand system presents several advantages. The structural derivation from consumer theory provides intuition for the interpretation of the underlying model parameters. This formal link to theory also enables the measurement of explicit structural metrics, such as consumer welfare. We demonstrate below instances in which welfare measurement provides valuable information to policy makers as well as marketers managing customer relationships. Finally, the logit demand system provides a parsimonious representation of complex substitution patterns in categories with differentiated products. Substitution patterns are explained by projecting consumer preferences onto a set of (observed or unobserved) product attributes. Thus, the parameter space is determined by the number of characteristics rather than by the number of products[1]. Berry (1994) provides a thorough discussion of the merits of the logit demand system versus alternative specifications for differentiated products.

In order to obtain managerial implications, the estimated demand system is combined with a model of competitive market structure. In most cases, the competitive model focuses on price-setting firms and the static Bertrand equilibrium concept. This approach recognizes that prices and sales are endogenously determined as an equilibrium in an environment in which consumers maximize utility subject to a budget constraint and

competing firms set prices strategically to maximize their profits. In the context of discrete choice demand systems, several authors have documented that failure to account for the endogeneity of strategic variables may generate estimation biases (e.g. Berry 1994) analogous to the well-known simultaneity problems of supply and demand systems.[2]

The remainder of this paper is structured as follows. In section 2, we discuss the derivation of the mixed multinomial logit demand system. In section 3, we discuss various methods for estimating the demand system using either individual choice data or aggregate data. In section 4, we outline several applications of these demand systems to address both policy and strategy concerns.

2. The Mixed Multinomial Logit Demand System

In this section, we describe the underlying consumer choice model generating the aggregate demand in a market. We use the mixed logit specification (McFadden and Train 2000), which adds normally-distributed random coefficients to the standard conditional logit choice model. One of the main advantages of this specification is parsimony. Consumer preferences are projected onto a set of exogenous product attributes, which greatly reduces the dimension of the estimation problem. For industries with a large number of alternatives, correlation in the consumer valuations of products is characterized by the underlying attributes. For example, this approach has been applied to aggregate data for automobiles (BLP 1995; Petrin 1999, and Sudhir 2001a), PCs (Bresnahan, Stern and Trajtenberg 1997), ready-to-eat cereals (Nevo 2001), airlines (Berry, Carnall and Spiller 1997), movie theaters (Davis 1997), Broadway Theater (Leslie 2001), movies (Moul 2001), retail gasoline (Manuszak 2000) and paper towels (Cohen 2000). It has also been used in the context of consumer level data (Goldberg 1995, and Fader and Hardie 1996). In many product markets, such product attributes may not be available or, the underlying attributes may be too intangible to measure accurately. In such instances, one can estimate a set of latent product attributes. Since packaged goods often differ primarily by intangible brand valuations, researchers estimate the joint distribution of brand preferences across households using a factor-analytic approach (Elrod 1988; Chintagunta 1994, and Elrod and Keane 1995 use factor models for consumer data and Chintagunta, Dubé and Singh 2001a,b—hereafter referred to as CDS—use store data). Similarly, Goettler and Shachar 2000 use a factor model for household choices of television viewing due to a lack of measurable attributes characterizing the various television programs.

Formally, we assume that for a market t ($t = 1, \ldots, T$), M_t consumers each select one of J brands in the category or opt for the no-purchase alternative, whose utility is normalized to 0. Each alternative j has attributes: (x_{jt}, ξ_{jt}). The vector x_{jt} includes product attributes and, potentially, exogenous marketing mix information. The vector, ξ_{jt} encompasses the effects of unobserved (to the econometrician) product attributes. In the next section, we discuss some of the issues that may arise due to these unmeasured attributes and how the literature has addressed them. Finally, the variable p_{jt} denotes alternative j's price in week t.

In a market t, the conditional utility consumer h derives from purchasing product j is given by:

$$u_{hjt} = \alpha_{hj} + x_{jt}\beta_h + \theta_h(Y_h - p_{jt}) + \xi_{jt} + \varepsilon_{hjt}.$$

The coefficients β_h capture consumer h's tastes for attributes, x. The parameter θ_h captures consumer h's marginal utility for income, Y_h. In the current context, income consists of the shopping budget for a trip[3]. The parameter α_{hj} captures household h's idiosyncratic perception of alternative j. The term ε_{hjt} is an i.i.d. mean-zero stochastic term capturing consumer h's idiosyncratic utility for alternative j in market t. In practice, the utilities, u_{hjt}, are unobserved (to the econometrician). Since ξ_{jt} accounts for randomness due to unmeasured product characteristics, the term ε_{hjt} reflects random variation in consumer choice behavior. In the current discussion, we assume that ε_{hjt} is drawn from a type I extreme value distribution, giving rise to the standard logit choice model. Previous work has also explored the use of correlated errors, such as the multivariate normal, giving rise to the probit choice model (McCullough and Rossi 1994). Below we discuss some of the limitations of an i.i.d. additive error.

Since we do not observe the true distribution of consumer preferences, we assume tastes, brand perceptions and the marginal utility of income are drawn from a multivariate normal distribution:

$$\begin{bmatrix} \beta_h \\ \theta \end{bmatrix} \sim N\left(\begin{bmatrix} \bar{\beta} \\ \bar{\theta} \end{bmatrix}, \Sigma \right)$$

where the vectors of means, $(\bar{\beta}, \bar{\theta})'$, and the covariance matrix, Σ, are parameters to be estimated.

In practice, the estimation of the matrix Σ could offset the potential parsimony-related benefits of the discrete choice model. Typically, researchers treat random tastes as i.i.d., so that only the diagonal elements need to be identified. Flexibility in substitution patterns are obtained by heterogeneity across consumers in their tastes for the product characteristics. However, as discussed above, when a vast set of product characteristics is not available, one may need to allow for correlation in the tastes themselves to capture rich substitution patterns. For instance, in the market structure literature (e.g., Elrod 1988), correlations are modeled for the product fixed-effects:

$$\alpha_h \sim N(\bar{\alpha}, \vartheta).$$

To alleviate the loss in degrees of freedom, a factor structure is then imposed on the covariance matrix, ϑ (see Elrod 1988 for an application using household data, and Chintagunta, Dubé and Singh 2002a for an application using aggregate data).

With individual choices, one can also model observed heterogeneity by interacting the product characteristics with individual-specific information, such as demographics. When products can be characterized by their geographic location, consumer heterogeneity can also be captured by the distance traveled (Capps et al. 2000). Using aggregate data, implementing the effects of consumer demographics has been achieved by sampling from the empirical distributions provided by the census bureau (Nevo 2001). A similar approach

has been used to sample from the empirical distribution of household locations for models of retail competition (Davis 1997; Manuszak 2000, and Thomadsen 2001).

As is now the convention in the literature, we simplify our notation by re-writing the consumer's indirect utility in terms of mean tastes and deviations from the mean:

$$u_{hjt} = \delta_{jt} + \mu_{hjt} + \varepsilon_{hjt}$$

where $\delta_{jt} = \bar{\alpha}_j + x_{jt}\bar{\beta} - \bar{\theta}p_{jt} + \xi_{jt}$ is common to all consumers and $\mu_{hjt} = x_{jt}\sigma v_h + L\omega_h$ is consumer-specific. σ is a vector of the square roots of the diagonal elements in Σ, L is the Cholesky decomposition of ϑ and v_h and ω_h are vectors of independent standard normals[4]. An advantage of this mixture of the normally-distributed random taste coefficients with the extreme value disturbance, is that we can integrate out the latter analytically. The probability q_{jt} that a consumer chooses a particular product j in week t has the following form:

$$q_{jt} = \int_{-\infty}^{\infty} \frac{\exp(\delta_{jt} + \mu_{hjt})}{1 + \sum_{i=1}^{J} \exp(\delta_{it} + \mu_{hit})} \phi(\lambda)\partial\lambda, \tag{1}$$

where $\lambda = (v, \omega)'$ and $\phi(\cdot)$ is the pdf of a standard normal. From the store manager's perspective, (1) represents the share of consumers entering the store in week t that purchase a unit of product j. Thus, the manager's expected demand for product j in store-week t is:

$$Q_{jt} = q_{jt}M_t. \tag{2}$$

The main motivation for using this random coefficients specification, as opposed to a simpler conditional logit (or homogeneous logit), is the need for flexible substitution patterns. Regardless of whether one estimates the parameters of the model using individual choice data or aggregate data, ignoring the effects of heterogeneity generates the conditional logit's restrictive substitution patterns at the market level. These unrealistic substitution patterns, in turn, restrict the equilibrium pricing behavior. For example, static Bertrand price-cost margins are increasing in the market share. Moreover, multiproduct firms are restricted to set a uniform margin for each of the products in their line (Besanko, Dubé and Gupta 2001, and Chintagunta 2001a). Accounting for heterogeneity, however, allows for flexible substitution patterns at the aggregate market level.

As discussed above, one could also use a correlated additive error, such as the probit model, which resolves the IIA problem at the individual level. The increasing use of the "mixed" logit versus a multinomial probit is mainly the relative ease of estimating the former versus the latter. In general, the probit strictly dominates the logit as it allows for freely-varying covariances (up to normality). The random coefficients probit also enables one to disentangle heterogeneity from simple non-IIA behavior at the consumer level. McFadden and Train (2000) show that the the mixed logit can be sufficiently flexible to approximate a broad set of parametric indirect utility functions, including the probit (see

Dalal and Klein 1988 for a related finding). In practice, the flexibility depends on the restrictions placed on the correlations in the random coefficients, ϑ. With aggregate data, the ability to integrate out the logit disturbance, as in (1), vastly increases the ease of implementation versus a probit. In fact, the aggregation of a probit may only be able to accommodate a small number of products (Chintagunta 2001b).

One of the attractive features of the discrete choice model is the ability to compute changes in consumer welfare in response to price or quality changes. A popular measure for welfare in these contexts is the Hicksian, or compensating, variation, which captures the dollar amount by which consumers would need to be compensated to maintain the same level of utility after the price or quality change. Typical applications include the measurement of the change in consumer welfare in response to a new product introduction (Trajtenberg 1989, and Petrin 1999) and mergers (Nevo 2000, and Dubé 2001). We denote an individual h's utility net of the extreme value taste shock as V_h and their marginal utility of income as θ_h. Suppose a policy is introduced that changes consumer valuations for each alternative from V_h^0 to V_h^1 (e.g. a change in the nature of competition causes prices to change). As derived in Small and Rosen (1981), assuming individual marginal value of income is held constant, individual h's change in welfare associated with this new policy is measured by:

$$CV_h = \frac{\log\left(\sum_{j=0}^J \exp(V_h^{zone})\right) - \log\left(\sum_{j=0}^J \exp(V_h^{chain})\right)}{\theta_h}. \tag{3}$$

Integrating across all the consumers provides the measure of aggregate change in consumer welfare:

$$\Delta W = \int CV_h \phi(v) \partial v. \tag{4}$$

3. Static Bertrand Oligopoly

The majority of this literature assume that firms set profit-maximizing prices in a static manner each period. This assumption is primarily based on tractability. In the applications section below, we check the validity of the assumption by testing for price deviations from the static Bertrand Nash equilibrium. The typical model assumes that in any given period t, a firm j with a product line $\Im_j$ sets prices to maximize its profits:

$$\prod_{jt} = \sum_{k\in\Im_j} (p_k - c_k)Q_k \tag{5}$$

where p_k denotes the price of good k, c_k denotes its per-unit cost (which is assumed to be constant), and Q_k denotes the demand for k as defined by (2). The maximization of (5) yields the following set of first-order conditions for firm j:

$$\sum_{l \in \ _j} (p_l - c_l) \frac{\partial Q_l}{\partial p_k} + Q_k = 0, \quad \text{for all } k \in \ _k, \tag{6}$$

From (6), the system of prices can be written in the following matrix notation:

$$\mathbf{p} = \mathbf{c} + \mathbf{\Phi}^{-1} \mathbf{Q} \tag{7}$$

where the matrix $\mathbf{\Phi}$ contains elements:

$$\mathbf{\Phi}_{ij} = -\frac{\partial Q_i}{\partial p_j}, \ \text{if } i, j \in \ _k \ \text{for some } k$$

$$= 0, \ \text{else.}$$

In some contexts, the model is amended to include both manufacturers and retailers to reflect the channel structure of retail goods. In a static framework in which both retailers and manufacturers set prices simultaneously, it can be shown (see Besanko, Gupta and Jain 1998) that equilibrium prices will have the following form:

$$\mathbf{p} = \mathbf{c} + \mathbf{\Phi}^{-1} \mathbf{Q} + \mathbf{\Theta}^{-1} \mathbf{Q} \tag{8}$$

where the matrix $\mathbf{\Theta}$ has elements $\Theta_{ij} = -(\partial Q_i / \partial p_j)$ for all i, j. This structure reflects the fact that all manufacturers sell through a common retailer. As a result, the observed shelf prices contain a double-marginalization—the retail margin and the wholesale margin. Researchers typically refer to the equilibrium pricing in this simultaneous-move channel model as the Vertical Nash equilibrium. To capture the fact that manufacturers/distributors may in practice commit to wholesale prices before retailers set their shelf prices, researchers may instead use the Stackelberg equilibrium concept. The Stackelberg equilibrium to the channel game also gives the double-marginalization outcome (see Sudhir 2001b for a comparison of these and other candidate static equilibria in similar channel games). These games assume very simple channel contracts in which wholesalers charge simple linear prices. Depending on the nature of the contracts written between retailers and wholesalers, one can observe many other forms of prices. For instance, if the wholesalers use two-part tariffs, one can obtain prices for which the double-marginalization is eliminated. In the application section below, we also look at a means of testing for deviations in prices from the static Nash equilibrium in a channel.

4. Estimation Issues

The econometric methodology for the estimation of choice models using either individual or aggregate data is well-documented. Rather than provide the technical details for estimation, we discuss some of the main issues that arise. For a detailed discussion of the econometric implementation of the individual mixed logit, see McFadden and Train

(2000). Empirically, this model specification has had a long history in marketing beginning with the finite mixture specification of Kamakura and Russell (1989) and the continuous mixture specification of Chintagunta, Jam and Vilcassim (1991). For a thorough discussion of the econometric implementation of the aggregate mixed logit, see BLP (1995). For the aggregate probit, see Chintagunta (2001b).

A primary concern in empirical papers using discrete choice models is the potential for estimation bias due to correlation between prices and the unobserved product attribute, ξ (e.g. Berry 1994, and BLP 1995). For instance, ξ enters $\boldsymbol{\Phi}$ in the margin equation (7); so prices will be correlated with the unobserved attributes. One solution to this problem is to include product fixed-effects to control for such unobservables. However, fixed-effects do not control for potential deviations from the mean effects of unobserved attributes in data sources containing multiple markets. Even after including a full set of product fixed-effects, several authors have still found evidence of biases in the price response parameter. For instance, Nevo (2001) finds such biases in aggregate data containing multiple time periods (quarters) and multiple geographic markets (Nielsen scantraks). Similarly, Besanko, Gupta and Jain (1998) document biases in weekly supermarket data. The analogous problem has also been documented in applications using individual choice data (Villas-Boas and Winer 1999; Villas-Boas and Zhao 2001, and Goolsbee and Petrin 2001). Typically, researchers using individual scanner data have assumed that the standard simultaneity problems associated with estimating supply and demand systems do not affect individual demand. In particular, individual choices are considered too small to shift aggregate demand and, thus, have an impact on pricing. The omitted variable bias, however, has been shown to generate biases in individual choice models. Some authors have devised limited information instrumental variables procedures that use additional exogenous covariates that are correlated with prices to resolve the endogeneity problem (see Nevo 2001 for market-level data, and Goolsbee and Petrin 2001 for consumer-level data). Alternatively, others have modeled the full data-generating process, including equilibrium price equations in addition to the demand equations (see BLP 1995 for market-level data, and Villas-Boas and Zhao 2001 for consumer-level data). The price endogeneity is solved by modeling the price equilibrium explicitly.

The key limitation of these instrumental variables procedures is the availability of suitable instruments. Valid instruments must be correlated with prices, but not with the omitted variables (e.g. unobserved attributes). Often, factor prices and exogenous production-related cost variables may be obtained from a number of sources, such as the Bureau of Labor Statistics. Such instruments are believed to shift marginal costs and, thus, correlate with prices. The ability to find viable cost-related instruments varies dramatically from one industry to another. Alternatively, when a broad scope of measurable product attributes are available, functions of these attributes can be used based on competitive assumptions (BLP 1995). When neither costs nor attribute combinations perform well as instruments, others have used information from other markets, such as prices. For instance, Nevo (2001) uses prices from other geographic markets. This approach assumes that unobserved attributes that are unaccounted for in the model and which correlate with prices are independent across markets. At the same time, the price co-movements across markets reflect common cost-related information.

5. Applications

The main advantage of using a structural derivation is the ability to measure economic variables explicitly, such as variable profits and consumer welfare. We now discuss several applications of choice models to measure the underlying economics of various industries. We summarize each of these applications in Table 1.

5.1. Identifying the Nature of Competition

Most of the strategic applications of the discrete choice demand systems require assumptions about the determination of prices, such as the static Bertrand Nash equilibrium (7). Moreover, identification of the models typically requires assumptions of independence across time and geographic space to generate enough "markets" for consistent estimation of the parameters. A natural research question involves tests for deviations from these simple forms of pricing. Using household scanner data for ketchup, Villas-Boas and Zhao (2001) model the supply-side using a vertical channel structure as a baseline model. To test the validity of the double-marginalization from assuming Stackelberg Nash equilibrium, they interact the implied retail margin in (8)[5] with the conduct parameter μ and the wholesale margins with conduct parameters $\theta_1, \ldots, \theta_J$ (one for each of the J products produced by each of the respective J manufacturers). Significance tests for these conduct parameters indicate whether equilibrium pricing embodies a single or a double marginalization. Similarly, Sudhir, Kadiyali and Chintagunta (2001) introduce conduct parameters into a Bertrand Oligopoly model, (7), for the US photographic film industry. They specify the conduct parameters as functions of

Table 1. Applications

	Data Type	Supply-Side Model	Objective
Identifying the Nature of Competition			
Villas-Boas and Zhao (2001)	Individual household-level ketchup purchases	Retail/wholesale channel	Test for nature of game
Sudhir, Kadiyali and Chintagunta (2001)	Quarterly national-level sales/prices of photo film	Repeated oligopoly	Test for presence of regime-switching
Bronnenberg and Mahajan (2001)	Weekly scantrac-level sales/prices of Tortilla Sauce	Not modeled	Test for spatial dependence
Strategy Simulation			
Thomadsen (2001)	Single day of restaurant-level sales/prices	Bertrand oligopoly	Measure market power— location vs. differentiation
Goettler and Shachar (2001)	1/4 hourly household-level TV channel choices	Static oligopoly (schedule)	Measure optimal schedule— assess rules-of-thumb
Goolsbee and Petrin (2001)	Individual household-level choice of broacast TV source	Bertrand oligopoly	Measure welfare gains from intro of Satellite TV
Chintagunta, Dubé and Singh (2002b)	weekly supermarket-level sales/prices various products	Static monopoly	Measure welfare impact of zone-pricing

exogenous state variables reflecting both shocks to demand and to production costs. As such, they capture the impact of exogenous shocks to the industry on the nature of market equilibrium and, thus, capture regime-switching. Finally, Bronnenberg and Mahajan (2001) provide counter-evidence to the standard assumption of the static model when applied to data for multiple geographic markets. Their spatial model reveals evidence of strong sources of dependence across geographic markets in both prices and market shares. These results put to question standard practices of pooling data from multiple geographic markets for estimation and, more recently, using prices from other markets as instruments (as described above).

5.2. Strategy Simulation

One of the main contributions of structural modeling is the ability to measure explicitly the underlying economic impact of an economic or strategic policy. The estimated structural parameters enable the researcher to characterize fully the underlying model of supply and demand in equilibrium. This model consisting of the system of equations (2) and (7) can, in turn, be used as a simulator of the industry. This simulator can be used to measure the impact of counter-factual "what if" scenarios on equilibrium prices and sales. For instance, Thomadsen (2001) estimates an equilibrium model using prices of hamburgers across fast-food restaurants in a metropolitan area. On the demand-side, he is able to distinguish between consumers' willingness-to-pay versus their willingness-to-travel. In measuring market power, he can disentangle the influence of geographic location and differentiation (e.g. brand) on prices and, thus, profits. Using the model, he then simulates the equilibrium prices under several hypothetical geographic configurations to illustrate the strong downward pressure on prices due to proximity of competitors. Similarly, Goettler and Shachar (2001) estimate demand using household panel data for television viewing. Using their results, they compute the static Nash equilibrium programming schedule. Comparing these results with the observed schedule, they are able to assess the industry's rule-of-thumb practices.

The ability to compute hypothetical prices also serves as a useful input for welfare measurement using (4). Goolsbee and Petrin (2001) estimate household-level demand for satellite, basic cable, premium cable and local antenna. To measure the welfare implications of the introduction of satellite television, they re-compute the static Bertrand equilibrium prices when satellite is removed from the choice set. As such, they can compare the total consumer welfare in dollars with and without satellite, providing a money metric for the gains to consumers from satellite. Similarly, CDS (2001b) estimate demand using weekly store-level data for several packaged goods product categories in a large Chicago supermarket chain. They assess the impact on consumer welfare and chain profitability of moving from chain-wide pricing to a more disaggregate zone-pricing strategy (stores are grouped into non-overlapping zones and prices are set independently for each zone). Computing the optimal prices and the resulting firm profits and consumer welfare under each policy, they can measure the impact of this pricing policy explicitly. Moreover, they are able to design more sophisticated store-level pricing policies that further enhance both store profits and consumer welfare.

6. Conclusions and Future Research Directions

A recent resurgence in the popularity of structural modeling has been, to a large extent, stimulated by advances in the ability to apply discrete choice demand systems to complex differentiated product categories while controlling for the endogeneity of prices. One of the main advantages of the structural approach is the ability to measure the economic impact of strategic policies with explicit economic metrics such as consumer welfare and firm profits. However, many challenges still exist in perfecting the models to provide accurate reflections of the true underlying data-generating process. Below, we summarize possible areas for future research.

A number of papers have now documented consistent evidence of endogeneity biases in the estimated effects from discrete choice models. As described in this paper, most of the evidence has been documented using aggregate data. However, as also described here, some recent studies show that analogous problems may occur with individual choice data. Often, individual data reflect a string of choices made by consumers over time, for example household scanner data. Estimating these individual choice models, while controlling for the types of endogeneity described herein typically requires additional modeling assumptions, especially in the context of such repeated choices at the consumer level. This literature would benefit with more investigation both of the sources of these biases as well as methods for controlling for these biases.

Most of the existing research focuses on prices as the exclusive strategic variable available to competing firms. However, many of the industries discussed above are also characterized by aggressive advertising. In some instances, location in either geographic space or even a more abstract attribute space was also shown to play an important role in determining the level of competition. A fruitful area for future research would involve the use of models that capture the fact that firms typically compete with a portfolio of strategic variables. Not only would such models provide a potentially more realistic portrayal of competitive behavior, they would also provide insight into the interaction between each of a firm's competitive tools.

In addition to recognizing the multiplicity of strategic variables available to a firm, research would also benefit from expanding the horizons beyond the simple static models. On the demand-side, a large body of marketing research has documented the long-term effects of strategic variables, especially advertising, on sales. In the context of structural models, researchers should consider the potential for carry-over effects of current strategic variables on demand. For instance, Erdem and Keane (1996) propose a model of consumer choice with learning effects, which could potentially influence firms' price-setting decisions. In general, if current prices have an impact on future demand, firms should have a forward-looking perspective when setting price levels. Clearly, demand-side dynamics imply that firm competition may also be dynamic in nature. Other sources of dynamics in competition could also arise from rigidities. In the vertical channel models discussed above, one may expect wholesalers to move at a much lower frequency (e.g., quarterly) than retailers (e.g. weekly). These rigidities would force wholesalers to have a forward-looking perspective when setting their wholesale prices. A small, but growing, literature on dynamic oligopoly should provide a vast new perspective for structural modelers (e.g. Pakes and McGuire 1994).

Several challenges still face researchers interested in structural modeling. Most difficult is the availability of suitable data for dealing with the endogeneity of strategic variables. While researchers increasingly find access to rich demand-side data (such as sales and prices collected through supermarket checkout scanners), similar supply-side data is often unavailable. As in Thomadsen (2001), one may even lack key covariates, such as sales, requiring careful modeling solutions to circumvent the data limitation. The discovery of rich new data sources will be key for the advance of this stream of research.

Finally, one of the standard criticisms of structural modeling is the difficulty in assessing the extent to which results are driven by the data versus driven by restrictive parametric assumptions. Several recent papers have begun to investigate the potential for more flexible semi and non-parametric solutions. Pinkse, Slade and Brett (2001) propose a semi-parametric model for estimating the matrix of cross-price responses with spatial competition. Brown and Matzkin (1996, 1998) propose a non-parametric approach for simultaneous equations. In the context of choice models, Berry and Pakes (2001) propose a pure hedonic model that relaxes the assumption of the Type I extreme value taste shock. This growing body of research should have a strong impact on structural modeling as increasingly flexible models are designed, without the limitations of potentially limiting parametric assumptions.

Notes

1. Note that a typical linear demand system for a category with J products require J^2 parameters just to capture the substitution patterns.
2. Several authors estimate the full equilibrium system, consisting of both demand and supply-side conditions. As we discuss in a subsequent section, one can also estimate the demand function separately using standard instrumental variables techniques.
3. In the following analysis, we do not address formally how households allocate total income to their shopping budgets.
4. Note that we remove the term $\theta_h Y_h$ from the equation as it will not be identified in the share equations below. This term drops out of the share equation as it is common to all the alternatives including no-purchase.
5. In fact, they use the retail margin derived from a Stackelberg equilibrium in which wholesalers move first. The price equations under this assumption are more sophisticated than those derived above. However, the main point is the test of double-marginalization.

References

Allenby, G. M. (1989). "A Unified Approach to Identifying, Estimating and Testing Demand Structures with Aggregate Scanner Data," *Marketing Science*, 8, 265–280.

Allenby, G. M., and P. E. Rossi. (1991). "There Is No Aggregation Bias: Why Macro Logit Models Work," *Journal of Business and Economic Statistics*, 9, 1–13.

Baker, J. B., and T. F. Bresnahan. (1985). "The Gains from Merger or Collusion in Product-Differentiated Industries," *The Journal of Industrial Economics*, 33, 427–444.

Berry, S. (1994). "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25, 242–262.

Berry, S., M. Carnall and P. T. Spiller. (1997). "Airline Hubs: Costs, Markups and the Implications of Customer Heterogeneity," Yale University, Working Paper.

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

Berry, S., and A. Pakes. (2001). "Estimating the Pure Hedonic Discrete Choice Model," Working Paper, Yale University.

Besanko, D., S. Gupta and D. Jain. (1998). "Logit Demand Estimation Under Competitive Pricing Behavior: An Equilibrium Framework," *Management Science*, November.

Besanko, D., J.-P. Dube, and S. Gupta. (2001). "Heterogeneity and Target Marketing Using Aggregate Retail Data: A Structural Approach," Working Paper, Kellogg Graduate School of Management.

Bresnahan, T. F., S. Stern, and M. Trajtenberg. (1997). "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *The Rand Journal of Economics*, 28, S17–S44.

Bronnenberg, B., and V. Mahajan. (2001). "Unobserved Retailer Behavior in Multi-Market Data: Joint Spatial Dependence in Market Shares and Promotion Variables," Working Paper, UCLA.

Brown, D. J., and R. Matzkin. (1996). "Testable Restrictions on the Equilibrium Manifold," *Econometrica*, 64, 1249–1262.

Brown, D. J., and R. Matzkin. (1998). "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand," Working Paper, Yale University.

Capps, C. et al. (2000). "Consumer Travel and Geographic Market Definition: An Application to Hospital Mergers," Working Paper, University of Illinois at Urbana-Champaign.

Chintagunta, P. (1994). "Heterogeneous Logit Model Implications for Brand Positioning," *Journal of Marketing Research*, 31, 304–311.

Chintagunta, P. (2001a). "Investigating Category Profit Maximization at a Retail Chain," *Journal of Marketing Research*, forthcoming.

Chintagunta, P. (2001b). "Endogeneity and Heterogeneity in a Probit Demand Model: Estimation Using Aggregate Data," *Marketing Science*, forthcoming.

Chintagunta, P., D. Jam, and N. Vilcassim. (1991). "Investigating Heterogeneity in Brand Preferences in Logit Models," *Journal of Marketing Research*, 28, 417–428.

Chintagunta, P., J. P. Dube, and V. Singh. (2002a). "Market Structure Across Stores: An Application of a Random Coefficients Model with Store Level Data," in P. H. Franses and A. L. Montgomery, eds., Econometric Models in Marketing (Advances in Econometrics, 16), JAI Press.

Chintagunta, P., J. P. Dube, and V. Singh. (2002b). "Balancing profitability and customer value: An application to zone-pricing by a supermarket chain," Working Paper, The University of Chicago Graduate School of Business.

Cohen, A. (2000). "Package Size and Discrimination in Paper Towels," Working Paper, University of Virginia.

Dalal, S. R., and R. W. Klein. (1988). "A Flexible Class of Discrete Choice Models," *Marketing Science*, 7, 232–235.

Davis, P. (1997). "Spatial Competition in Retail Markets," Yale University, Working Paper.

Dubé, J. P. (2001). "Product Differentiation and Mergers in the Carbonated Soft Drink Industry," Working Paper, University of Chicago Graduate School of Business.

Elrod, T. (1988). "Choice Map: Inferring a product market map from panel data," *Marketing Science*, 7, 21–40.

Elrod, T., and M. P. Keane. (1995). "A factor-analytic probit model for representing the market structure in panel data," *Journal of Marketing Research*, 32, 1–16.

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.

Fader, P. S., and B. G. S. Hardie. (1996). "Modeling Consumer Choice Among SKUs," *Journal of Marketing Research*, 33, 442–452.

Goettler, R., and R. Shachar. (2001). "Spatial Competition in the Network Television Industry," Working Paper, Carnegie Mellon.

Goldberg, P. K. (1995). "Product Differentiation and Oligopoly in International Markets: The case of the U.S. Automobile Industry," *Econometrica*, 63, 891–951.

Goolsbee, A., and A. Petrin. (2001). "Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," Working Paper, The University of Chicago GSB.

Guadagni, P. M. and J. D. C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data." Marketing Science, 2, 203–238.

Hausman, J. A. (1997). "Valuation of New Goods Under Perfect and Imperfect Competition." In T. Bresnahan, and R. J. Gordon (eds.), *The Economics of New Goods*, Chicago: University of Chicago Press.

Kadiyali, V., K. Sudhir, and V. R. Rao. (2000). "Structural Analysis of Competitive Behavior," *International Journal of Research in Marketing*, 18, 161–485.

Kamakura, W. A., and G. J. Russell. (1989). "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26, 379–390.

Leslie, P. J. (2001). "Price Discrimination in Broadway Theatre," Working Paper, UCLA.

Manuszak, M. D. (2000). "Firm Conduct and Product Differentiation in the Hawaiian Retail Gasoline Industry," Working Paper, Carnegie Mellon University.

McCullough, G., and P. E. Rossi. (1994). "An exact likelihood analysis of the multinomial probit model," *Journal of Econometrics*, 64, 207.

McFadden, D. (1973). "Conditional Logit Analysis of Discrete Choice Behavior." In P. Zarembka (ed.), *Frontiers of Econometrics*, New York: Academic Press.

McFadden, D., and K. Train. (2000). "Mixed MNL Models for Discrete Response," *Journal of Applied Economerics*, 15, 447–470.

Moul, C. (2001). "Saturation and the Demand for Motion Pictures," Working Paper, Washington University St. Louis.

Nevo, A. (2001). "Measuring Market Power in the Ready-To-Eat Cereal Industry," *Econometrica*, 69, 307–340.

Nevo, A. (2000). "Mergers and Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *The Rand Journal of Economics*, 31, 395–421.

Pakes, A., and P. McGuire. (1994). "Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model," *The Rand Journal of Economics*, 79, 555–589.

Petrin, A. (1999). "Quantifying the Benefits of New Products: The Case of the Mini-van," Working Paper, University of Chicago.

Pinkse, J., M. E. Slade, and C. Brett. (2001). "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, forthcoming.

Small, K. A., and H. S. Rosen. (1981). "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, 49, 105–130.

Sudhir, K. (2001a). "Competitive Pricing Behavior in the Auto Market: A Structural Analysis," *Marketing Science*, 20, 42–60.

Sudhir, K. (2001b). "Structural Analysis of Competitive Pricing in the Presence of a Strategic Retailer," *Marketing Science*, 20, 244.

Sudhir, K., P. Chintagunta, and V. Kadiyali. (2001). "Investigating Patterns of Time-Varying Competitive Behavior," Working Paper, Yale School of Management.

Thomadsen, R. (2001). "Price Competition in Industries with Geographic Differentiation: The Case of the Hamburger Segment of the Fast Food Industry," Working Paper, Columbia University.

Trajtenberg, M. (1989). "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *The Journal of Political Economy*, 87, 444–479.

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

Villas-Boas, M., and Y. Zhao. (2001). "The Ketchup Marketplace: Retailer, Manufacturers, and Individual Consumers," Working Paper, UC Berkeley.