Competitive Price Discrimination Strategies in a Vertical Channel Using Aggregate Retail Data

David Besanko • Jean-Pierre Dubé • Sachin Gupta

Kellogg School of Management, Northwestern University, Evanston, Illinois 60208
Graduate School of Business, University of Chicago, Chicago, Illinois 60637
Johnson Graduate School of Management, Cornell University, Ithaca, New York 14853
d-besanko@kellogg.northwestern.edu • jdube@gsb.uchicago.edu • sg248@cornell.edu

We explore opportunities for targeted pricing for a retailer that only tracks weekly store-level aggregate sales and marketing-mix information. We show that it is possible, using these data, to recover essential features of the underlying distribution of consumer willingness to pay. Knowledge of this distribution may enable the retailer to generate additional profits from targeting by using choice information at the checkout counter. In estimating demand we incorporate a supply-side model of the distribution channel that captures important features of competitive price-setting behavior of firms. This latter aspect helps us control for the potential endogeneity generated by unmeasured product characteristics in aggregate data. The channel controls for competitive aspects both between manufacturers and between manufacturers and a retailer. Despite this competition, we find that targeted pricing need not generate the prisoner's dilemma in our data. This contrasts with the findings of theoretical models due to the flexibility of the empirical model of demand. The demand system we estimate captures richer forms of product differentiation, both vertical and horizontal, as well as a more flexible distribution of consumer heterogeneity.
(*Price Discrimination; Competition; Channels of Distribution; Scanner Data*)

1. Introduction

There is strong evidence in the marketing literature that consumers are very different in terms of their willingness to pay for products. Empirical studies report high variability in consumers' responsiveness to marketing-mix variables such as prices, in-store displays, and feature advertisements, as well as in their intrinsic preferences for brands. This empirical consistency suggests that firms have an opportunity to price discriminate profitably rather than charge a uniform price to all consumers. There is a vast literature in economics on the theory of price discrimination (see, for example, Varian 1989 for a review of this literature), and a number of important papers

in marketing (e.g., Moorthy 1984, Narasimhan 1984) have discussed different forms of price discrimination and how they might be implemented in practice. Rossi et al. (1996) demonstrate the benefits of targeted picing using household scanner panel data.

While price discrimination by a monopolist always leads to profits that are at least as large as those under uniform pricing, the competitive implications of price discrimination in oligopoly markets are more subtle. Shaffer and Zhang (1995), for example, show in a theoretical model that targeted couponing leads to a prisoner's dilemma in which all manufacturers issue coupons without profitably increasing their prices. Chen et al. (2001) find that competing firms may gain incremental profits when consumer targetability

is very imprecise. However, as targetability improves, they also find the prisoner's dilemma. Corts (1998) finds that price discrimination by all competing firms may lead to "all-out competition," which results in prices and profits getting reduced in all market segments. However, Corts (1998) also provides counterexamples for which competitive price discrimination may lead to increased profits. Hence, in oligopoly markets it is an *empirical* question whether or not the particular market demand conditions are favorable for price discrimination.

To practice price discrimination successfully, firms need detailed knowledge of the distribution of consumers' willingness to pay. In the marketing science literature, numerous methodologies have been developed to estimate discrete choice models of consumer demand and to characterize distributions of consumer heterogeneity (see Allenby and Rossi 1999 for a review). With few exceptions (e.g., Zenor and Srivastava 1993, Chintagunta 2002), the tradition in the marketing literature has been to use household panel data to model unobserved consumer heterogeneity (e.g., Kamakura and Russell 1989, Chintagunta et al. 1991). However, in many contexts the appropriate microdata may be unavailable. Typically, sample sizes in scanner panel data are small at the level of an individual store, complicating the estimation of a store-level demand system. Some marketing researchers have questioned the representativeness of purchase behavior of panelist households (Bucklin and Gupta 1999, Gupta et al. 1996). In these instances, store-level data may still contain useful information for econometric analyses of consumer demand. Aggregate store-level data are more widely and easily available, especially to retailers. We also believe that it is an interesting academic question as to whether consumer heterogeneity can be recovered in the kinds of aggregate store-level data that are typically available to retailers.

In this paper, we use aggregate weekly store-level data to study price discrimination strategies. The context for our work is a vertical channel consisting of competing manufacturers who sell through a common retailer. We estimate a structural model of manufacturer and retailer pricing in this channel, along with a discrete choice model of demand. Consumer heterogeneity is modeled as a finite number of discrete segments. We assume that the firm, manufacturer, or retailer can use the estimated demand system, combined with a single observed transaction at the checkout counter, to compute a set of posterior segment probabilities conditional on the current transaction. The firm can then offer customized coupons based on the current purchase for redemption on a future purchase occasion to implement a discriminatory pricing scheme. We show that such targeted pricing need not lead to a prisoner's dilemma. Because we model the channel structure, we are also able to contrast couponing by the retailer with couponing by the competing manufacturers.

We use a discrete representation of consumer heterogeneity (Berry et al. 1997). Specifically, consumers are assumed to belong to a finite number of latent classes that differ in terms of their brand preference and marketing-mix responsiveness parameters. The membership of consumers in latent classes is probabilistic and to be determined from the data. This approach has been termed a finite mixture or latent-class approach (Kamakura and Russell 1989, Wedel and Kamakura 1998). Such models have received considerable attention from practitioners as well, because the latent classes correspond closely with managers' notion of market segments. In this respect, we expect the econometric model to correspond to the way managers view demand when they determine prices. Our approach does not rely heavily on parametric assumptions to recover the underlying distribution of heterogeneity. Thus, we are less concerned that our results are driven by restrictive assumptions on the model primitives. When applied to aggregate data, latent-class models are computationally more tractable relative to continuous models of heterogeneity (e.g., Berry et al. 1995) because they do not require the simulation of high-dimensional integrals. Finally, managers may only be able to address a finite number of market segments when designing pricing or promotion strategies. In these situations, finite mixture models offer a managerially appealing and useful representation of the marketplace.

Our model provides several useful insights. We use the model to characterize the underlying consumer segments and their relative sensitivities to the marketing mix. The channel model allows us to control for competitive reactions of channel members

to targeting initiatives. We then illustrate the potential increase in retailer profits that could arise if the retailer was able to engage in price discrimination by offering segment-specific prices through customized coupons. With point-of-sale data, the retailer has only one observation for each consumer, hence classification of consumers to segments is imperfect. Even with imperfect classification, we find that the increase in retail profit relative to optimal uniform pricing can be as high as 11%. We also explore the incentives of manufacturers to engineer discriminatory retail prices. In our data from the ketchup market, we find that issuing coupons is a dominant strategy for both manufacturers, and both manufacturers are strictly better off.

We organize this paper as follows. In §2, we develop our models of demand and pricing. In §3, we describe the econometric procedure used to estimate the model. In §4, we present demand estimation results using data on the ketchup market. In §5, we demonstrate how segment-specific pricing strategies can improve the profitability of the retailer and of manufacturers. In §6, we summarize and conclude with a discussion of potential limitations and extensions.

2. Model

2.1. Utility and Demand

As stated in the introduction, the estimation procedure that we employ uses aggregate store-level sales data. We now derive an aggregate demand system built on the assumption that the underlying consumer behavior derives from a discrete choice framework (McFadden 1981).

We assume the potential market size for the category in any store-week to be the number of shopping trips to that store in that week. We also assume that the consumer's decisions of when to shop and at which store to shop are independent of prices in the category. On each shopping occasion, each consumer chooses whether or not to purchase in the category of interest (the purchase incidence decision),[1]

---

[1] Alternatively, one could model the time between category purchases. Wheat and Morrison (1990) compare these two alternative approaches and argue that modeling purchase incidence is almost always preferable to modeling interpurchase time.

and which product to purchase. Formally, we assume that on a given shopping trip in week t ($t = 1, \ldots, T$) consumers select one of J brands in the category (with a typical product indexed by i or j) or opt for the no-purchase alternative (also called the "outside good," indexed by 0). In week t, each brand j has three attributes: $(x_{jt}, p_{jt}, \xi_{jt})$. The vector x includes brand-specific indicator variables (thus allowing for the brand-specific constants) as well as marketing-mix variables (e.g., feature or display). The variable p denotes the brand's shelf price. Finally, ξ encompasses the effects of unobserved (to the econometrician) product attributes.

The underlying utility function in our model is quasi-linear for the products inside the category as well as the outside good, which is a numeraire representing expenditures on all items in the store outside the category (see Chintagunta 1993 for a detailed treatment). Thus, nonpurchase in the category of interest corresponds to allocating the entire shopping budget to this numeraire. For a shopping trip during week t, the indirect utility that consumer h ($h = 1, \ldots, H$) derives from purchasing product j in the category is given by

$$u_{hjt} = x_{jt}\beta_h - \alpha_h p_{jt} + \xi_{jt} + \omega_{hjt}. \tag{1}$$

The coefficients (β_h, α_h) capture consumer h's brand preferences and responsiveness to marketing variables, x, and price, p. The term ω_{hjt} is an i.i.d. mean-zero stochastic term capturing consumer h's idiosyncratic utility for alternative j during week t. The mean indirect utility of not purchasing is simply the marginal utility of income times the total shopping budget. Because this term enters the utilities of each of the inside goods as well as the outside good, it is not identified in the estimation process. Consequently we drop it, effectively normalizing the indirect utility of the outside good to be mean zero with the same i.i.d. stochastic component ω_{h0t}. By assuming that ω_{hjt} is drawn from a type I extreme value distribution, we obtain the multinomial logit model of choice (McFadden 1981).

A possible critique of our formulation is that at the individual level the model suffers from the independence of irrelevant alternatives (IIA) restriction between the products in the category

and the no-purchase alternative. Consequently, we also examined a non-IIA nested logit specification with a purchase/no-purchase nest and found that the predictive fit of this model in our data was inferior to the multinomial logit model (these results are reported in an appendix available at http://mansci.pubs.informs.org/ecompanion.html).

As mentioned above, the term ξ_{jt} captures a product attribute that is observed by the consumer, but not by the econometrician. Examples of such variables are exposure to television advertising, product availability at retail, and shelf space allocations. These variables not only affect brand choices of consumers, but may also influence price setting by firms. For example, national advertising by manufacturers, which can enhance brand salience and image, has been shown to be positively correlated with prices (Ackerberg 2001). Similarly, shelf space allocations, which vary considerably across stores, have been shown to have an important effect on consumer demand (Drèze et al. 1994). If this attribute is observed by firms, then it will influence the firms' pricing decisions, generating a simultaneity bias (Berry et al. 1995, Besanko et al. 1998).

This specification assumes consumer-specific preference and responsiveness vectors (β_h, α_h). To model consumer heterogeneity, we adopt a random coefficients approach. We use a discrete approximation to the parameter distribution, an aggregate analogue to the latent-class models used for household purchase data (Kamakura and Russell 1989). In particular, we assume that consumers belong to one of K segments, where each segment k is characterized by its own parameter vector, (α^k, β^k).

We denote the size of segment k as λ^k, which may be interpreted as the fraction of shopping trips arising from segment k. Note that when we only observe aggregate data, we cannot identify the proportion of consumers who belong to a given segment. We can only identify the proportion of shopping trips that arise from a given segment. Within a segment k, the share of total trips for brand j, S_{jt}^k, is simply the probability that consumers of type k purchase brand j and is given by

$$S_{jt}^k = \frac{\exp(x_{jt}\beta^k - \alpha^k p_{jt} + \xi_{jt})}{1 + \sum_{i=1}^{J} \exp(x_{it}\beta^k - \alpha^k p_{it} + \xi_{it})}.$$

Integrating across segments, we obtain the market share S_{jt} for product j in a given week t:[2]

$$S_{jt} = \sum_{k=1}^{K} \lambda^k \left(\frac{\exp(x_{jt}\beta^k - \alpha^k p_{jt} + \xi_{jt})}{1 + \sum_{i=1}^{J} \exp(x_{it}\beta^k - \alpha^k p_{it} + \xi_{it})} \right). \quad (2)$$

This aggregate demand system avoids the problems associated with IIA, an artifact of the i.i.d. assumption for ω_{hjt}. By mixing over the latent segments, we obtain more flexible aggregate substitution patterns than in the case of the homogeneous logit.

2.2. The Channel Structure
We now postulate the vertical channel model that describes the strategic generation of equilibrium retail shelf prices. In our model, oligopolistic manufacturers set wholesale prices and sell through a monopoly retailer. Assuming noncooperative behavior, the product shelf prices exhibit a double marginalization, including both a wholesale margin and a retail margin. Our baseline model specification corresponds with what Choi (1991) calls the Manufacturer Stackelberg (MS) game (Berto Villas-Boas 2002).[3] The key elements of the model are as follows. First, the retailer acts as a monopolist in its local area, where the size of this local market is M.[4] Second, the retailer sets a uniform price for all K segments. Third, there are N manufacturers, with a typical manufacturer denoted by n. Each manufacturer offers a set B_n of brands, with $\bigcup_{n=1}^{N} B_n = J$. Finally, the game between manufacturers and the retail chain unfolds in two stages. In the first stage, each manufacturer n takes the wholesale prices of other manufacturers as given, and chooses its set of wholesale prices, $\{w_j, j \in B_n\}$, to maximize its product line profits. In the second stage, the retailer takes wholesale prices as given, and chooses the set of

---

[2] The use of aggregate data prevents us from estimating consumer stockpiling or purchase acceleration behaviors in our static model. The possible implications of this are discussed in §5.5.

[3] We also compared the MS game with alternative models in terms of predictive fit and implied retail margins and found that it provides more reasonable estimates. Details are available from the authors.

[4] This assumption is broadly consistent with retailer conventional wisdom that most consumers shop at the same store week after week, often the one closest to their home or workplace (Slade 1995, Besanko et al. 1998).

retail prices $\{p_j, j \in J\}$ to maximize its overall category profits.

In the remainder of this section, we derive the subgame perfect Nash equilibrium in prices to this channel game. We begin with the second stage of the game, deriving the optimal prices of the retailer. Then, we move to the initial stage of the game and derive the optimal wholesale prices.

2.2.1. Retailer. The retailer takes the wholesale prices as given, and acts as a monopolist in pricing the whole category. The retailer's category management problem is

$$\max_{p_1, \dots, p_J} \Pi_R = \sum_{i=1}^{J} (p_i - w_i) \sum_{k=1}^{K} \lambda^k S_i^k M.$$

The first-order condition for a typical brand j is

$$\sum_{i=1}^{J} (p_i - w_i) \left(\sum_{k=1}^{K} \lambda^k \frac{\partial S_i^k}{\partial p_j} M \right) + \sum_{k=1}^{K} \lambda^k S_j^k M = 0.$$

We can rewrite the system of first-order conditions for brands $1, \dots, J$ in matrix form as:[5]

$$\mathbf{\Omega}(\mathbf{p} - \mathbf{w}) + \mathbf{v} = \mathbf{0}, \quad (3)$$

where

$$\mathbf{p} \equiv \begin{bmatrix} p_1 \\ \vdots \\ p_J \end{bmatrix}_{J \times 1}, \quad \mathbf{w} \equiv \begin{bmatrix} w_1 \\ \vdots \\ w_J \end{bmatrix}_{J \times 1},$$

$$\mathbf{v} \equiv \begin{bmatrix} \sum_{k=1}^{K} \lambda^k S_1^k M \\ \vdots \\ \sum_{k=1}^{K} \lambda^k S_J^k M \end{bmatrix}_{J \times 1}, \quad (4)$$

and

$$\mathbf{\Omega} \equiv \begin{bmatrix} -\sum_{k=1}^{K} \lambda^k \alpha^k S_1^k (1 - S_1^k) M & \cdots & \sum_{k=1}^{K} \lambda^k \alpha^k S_1^k S_J^k M \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{K} \lambda^k \alpha^k S_J^k S_1^k M & \cdots & -\sum_{k=1}^{K} \lambda^k \alpha^k S_J^k (1 - S_J^k) M \end{bmatrix}_{J \times J}.$$

$$(5)$$

[5] Note that $\partial S_j^k / \partial p_j = -\alpha^k S_j^k (1 - S_j^k)$ and $\partial S_i^k / \partial p_j = \alpha^k S_i^k S_j^k$.

The optimal retail prices highlight another key implication of modeling consumer heterogeneity. It is straightforward to show that if consumers were assumed to be homogeneous, the equilibrium retail prices would reflect equal markups over wholesale prices. By contrast, a model with consumer heterogeneity allows the equilibrium retail markups to vary from brand to brand within the category. One can also show that the homogeneous logit model implies that the rate at which the retailer passes through changes in the wholesale prices to the retail prices would be driven entirely by market shares. Furthermore, the cross-pass-through rates would be strictly negative (Sudhir 2001). In contrast, the proposed heterogeneous logit model is sufficiently flexible to allow for a wide range of pass-through rates, which is consistent with empirical evidence (e.g., Besanko et al. 2003).

2.2.2. Manufacturer. In the MS channel game, we assume that manufacturers move first, enabling them to condition on the anticipated competitive response of retailers to posted wholesale prices. The profit-maximization problem of manufacturer n is

$$\max_{\{w_i, i \in B_n\}} \Pi_n = \sum_{i \in B_n} (w_i - mc_i) \sum_{k=1}^{K} \lambda^k S_i^k M,$$

where mc_i is the manufacturer's marginal cost for brand i. The first-order condition for a typical brand j is

$$\sum_{i=1}^{J} (w_i - mc_i) \Upsilon_{ji} \left(\sum_{k=1}^{K} \lambda^k \sum_{l=1}^{J} \frac{\partial S_i^k}{\partial p_l} \frac{\partial p_l}{\partial w_j} M \right) + \sum_{k=1}^{K} \lambda^k S_j^k M = 0, \quad (6)$$

$$\Upsilon_{ji} = \begin{cases} 1 & \text{if brands } j \text{ and } i \text{ are offered} \\ & \text{by the same manufacturer,} \\ 0 & \text{if brands } j \text{ and } i \text{ are offered} \\ & \text{by different manufacturers.} \end{cases} \quad (7)$$

Unlike the retailer, who takes manufacturers' prices as given, the manufacturers explicitly account for the retailer's reactions to wholesale prices, $\partial p_l / \partial w_j$. To

compute the system of retail pass-through rates, we totally differentiate the first-order conditions (3):

$$\Delta \begin{bmatrix} \dfrac{\partial p_1}{\partial w_j} \\ \vdots \\ \dfrac{\partial p_J}{\partial w_j} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial S_j}{\partial p_1} \\ \vdots \\ \dfrac{\partial S_j}{\partial p_J} \end{bmatrix}$$

where

$$\Delta = \begin{bmatrix} 2\dfrac{\partial S_1}{\partial p_1} + \sum\limits_{l=1}^{J}(p_l - w_l)\dfrac{\partial^2 S_l}{\partial p_1^2} & \cdots & \dfrac{\partial S_J}{\partial p_1} + \dfrac{\partial S_1}{\partial p_J} + \sum\limits_{l=1}^{J}(p_l - w_l)\dfrac{\partial^2 S_l}{\partial p_J \partial p_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial S_1}{\partial p_J} + \dfrac{\partial S_J}{\partial p_1} + \sum\limits_{l=1}^{J}(p_l - w_l)\dfrac{\partial^2 S_l}{\partial p_1 \partial p_J} & \cdots & 2\dfrac{\partial S_J}{\partial p_J} + \sum\limits_{l=1}^{J}(p_l - w_l)\dfrac{\partial^2 S_l}{\partial p_J^2} \end{bmatrix}_{J \times J}. \tag{8}$$

Using this result, we can rewrite the system of first-order conditions for brands $1, \ldots, J$ in matrix form as

$$\Omega \Delta^{-1} \Phi (\mathbf{w} - \mathbf{mc}) + \mathbf{v} = \mathbf{0}, \tag{9}$$

where Ω and $\mathbf{v}$ are defined as in (4) and (5), respectively, $\mathbf{mc} = (mc_1, \ldots, mc_J)'$ and

$$\Phi \equiv \begin{bmatrix} -\Upsilon_{11}\left(\sum\limits_{k=1}^{K}\lambda^k \alpha^k S_1^k(1-S_1^k)M\right) & \cdots & \Upsilon_{1J}\left(\sum\limits_{k=1}^{K}\lambda^k \alpha^k S_1^k S_J^k M\right) \\ \vdots & \ddots & \vdots \\ \Upsilon_{J1}\left(\sum\limits_{k=1}^{K}\lambda^k \alpha^k S_J^k S_1^k M\right) & \cdots & -\Upsilon_{JJ}\left(\sum\limits_{k=1}^{K}\lambda^k \alpha^k S_J^k(1-S_J^k)\lambda^k M\right) \end{bmatrix}_{J \times J}.$$

The full vertical equilibrium of the model is characterized by the retailer's J first-order conditions, (3), the manufacturers' J first-order conditions, (9), and the J demand equations in market-share form, (2). To ensure that our estimates constitute a valid equilibrium, in the empirical analysis we check that our results satisfy the second-order sufficient conditions of the model. For the coupon analysis, because the equilibria are obtained numerically, we also check whether the algorithm converges to the same point from multiple starting values. Finally, while we do not observe the wholesale prices in our data, we are able to characterize the equilibrium shelf prices as

$$\mathbf{p} = \mathbf{mc} - (\Omega \Delta^{-1} \Phi)^{-1} \mathbf{v} - \Omega^{-1} \mathbf{v}. \tag{10}$$

3. Estimation

We now outline the estimation procedure for the parameters of the equilibrium model developed in the previous section. Defining $X_{jt} \equiv [x_{jt}, p_{jt}]$ and $\hat{\beta}^k \equiv (\beta^k, \alpha^k)$, we let $\delta_{jt} \equiv X_{jt}\hat{\beta}^1 + \xi_{jt}$ denote segment 1's mean utility for product j, and $\delta_t \equiv (\delta_{1t}, \ldots, \delta_{Jt})$ denote the entire vector of segment 1's mean utilities. Furthermore, let $\hat{\beta}^{*k} \equiv \hat{\beta}^k - \hat{\beta}^1$, $k = 2, \ldots, K$ denote the difference in parameters relative to segment 1. Using this notation, we can rewrite the share equations in (1) as

$$\overline{S}_j(X_t, \delta_t; \Theta) = \lambda^1 \frac{\exp(\delta_{jt})}{1 + \sum_{i=1}^{J}\exp(\delta_{it})} + \sum_{k=2}^{K}\lambda^k \frac{\exp(X_{jt}\hat{\beta}^{*k} + \delta_{jt})}{1 + \sum_{i=1}^{J}\exp(X_{it}\hat{\beta}^{*k} + \delta_{it})}, \tag{11}$$

where $\Theta \equiv (\hat{\beta}^{*2}, \ldots, \hat{\beta}^{*K}, \lambda^1, \ldots \lambda^K)$ denotes the full parameter vector to be estimated, $X_t \equiv (X_{1t}, \ldots, X_{Jt})$ denotes the vector of product characteristics across brands, and $\overline{S}_j(\cdot, \cdot; \cdot)$ denotes the market share *function* for brand j.

On the demand side, we assume $E(x_{jt}\xi_{jt} \mid x_{jt}) = 0$ and $E(p_{jt}\xi_{jt} \mid p_{jt}) \neq 0$. Following Berry (1994), we then invert (11) to recover the vector $\delta_t(\Theta)$ of mean utilities of segment 1 as a function of parameter vectors Θ and set up the estimation procedure in terms of δ_t. Because the inverse of (11) cannot be obtained analytically, we use the contraction mapping of Berry et al. (1995). The advantage of using δ_{jt} for estimation is that the prediction error, $\delta_{jt} - X_{jt}\hat{\beta}^1$, is simply the unobserved product characteristic, ξ_{jt}. The fact that ξ_{jt} enters $\delta_{jt}(\Theta) \equiv X_{jt}\hat{\beta}^1 + \xi_{jt}$ linearly allows us to use standard instrumental variable methods to control for the potential correlation between p_{jt} and ξ_{jt}.

On the supply side, we assume that mc_{jt} is explained by factor prices, c_t, and a random component (possibly unobserved factor prices): $mc_{jt} = c_t \gamma_j + \eta_{jt}$, where we assume $E(c\eta_j \mid c) = 0$ (the unobserved factor prices of production are conditionally independent of the observed factor prices). Our linear specification implicitly assumes a fixed-proportions (i.e., Leontief) production technology. For products such as ketchup, this assumption may not be unreasonable, at least in the short run. Substituting the marginal cost specification into equilibrium retail shelf-price equations (10) yields the pricing equations that we estimate. To be consistent with cost-minimizing input choices by manufacturers, the specification of mc_j should be homogeneous of degree one. However, we do not observe all factors of production, and for this reason we also include product-specific intercepts in mc_j to improve fit.

We now set up the generalized method of moments (GMM) procedure to estimate the system of price and demand equations. We outline the technical details of GMM below, referring the interested reader to Hansen (1982) for a formal discussion. We construct moments of the data-generating process based on the structural model; these take the form of orthogonality conditions. Parameters are estimated by making the sample analogue of the orthogonality conditions as close to zero as possible. Let $\varepsilon_t = \begin{bmatrix} \eta_t \\ \xi_t \end{bmatrix}$ be a $(2J \times 1)$ matrix with the prediction error for marginal costs and the unobserved attributes for each of the products in store-week t. We can construct orthogonality conditions using any set of covariates, Z_t, that are mean independent of ε_t. To construct this I-dimensional vector of instruments, Z_t, we include the exogenous product characteristics as well as factor costs. Formally, our data-generating process is described by the conditional mean independence assumptions, $E(\varepsilon_t \otimes Z_t \mid Z_t) = 0$, and the covariance restrictions $E(\varepsilon_t \varepsilon_t' \mid Z_t) = \mathbf{\Omega}$ a finite $(2J \times 2J)$ matrix and $E(\varepsilon_t \varepsilon_\tau') = 0$, $t \neq \tau$. We fit these moments to our data using the efficient GMM procedure of Hansen (1982).

The basic intuition for the identification of segments in aggregate data depends on asymmetric substitution patterns in the observed market shares. In effect, we fit heterogeneity using observed non-IIA aggregate substitution patterns. In the Appendix, we illustrate the identification using a simulation experiment with six brands and five segments. Note that we determined the number of mass points by adding segments until one of the segment sizes was not statistically different from zero.

4. Demand Estimation Results Using Ketchup Data

4.1. Data

We apply the model to a sample of weekly store-level data for ketchup. We use data on prices, market shares, in-store displays, and feature activity for the four largest products in the category. The data are collected by the ACNielsen Company in Springfield, Missouri, using store checkout scanners. We focus on nine stores belonging to a single chain during a 102-week period between 1986 and 1988.

To compute the (unconditional) brand market shares, we divide the total unit sales of each brand by an estimate of the total number of store trips in the given week, and the share of the no-purchase alternative is then one minus the sum of the brand market shares. Unlike many other store datasets, in our data we do not observe the total weekly store traffic, so we infer the total number of trips by using a panel of 2,500 households. Our estimate of store traffic is obtained by projecting the number of panelist trips to each store to the total population. With respect to the other variables, price is measured as the retail shelf price per ounce, net of in-store promotional price cuts. The feature variable is a 0–1 indicator for whether the product was advertised in a newspaper. Similarly, the display variable is a 0–1 indicator for whether the product appeared on special display in the store.

On the supply side, we use factor prices for labor and materials costs collected by the Bureau of Labor Statistics. Labor costs for ketchup consist of the average hourly earnings of production workers in the canned fruits and vegetables industries. For materials costs, we use the price index for tomatoes and the price index for glass containers. Because these data are reported on a monthly frequency, we use the lin-

Table 1 Descriptive Statistics for Ketchup Data

	Brand shares (conditional)	Price ($/oz)	Feature (% store-weeks)	Display (% store-weeks)
Heinz 32	0.394	0.043	9.7	4.9
Hunts 32	0.211	0.042	1.2	1.7
Heinz 28	0.314	0.051	8.2	2.7
Heinz 44	0.081	0.047	0	0

Factor	Price index
Labor costs	8.26
Tomatoes	98.93
Glass containers	111.15

ear filtering process suggested by Slade (1995) to convert the monthly data to weekly data.[6]

The ketchup data include three sizes of the leading national brand, Heinz, and one size of the second largest competitor, Hunts. These products account for over 70% of category volume. We report summary statistics for the prices, display, and feature activities of each of these products along with the factor prices in Table 1. We can see that Heinz is the market share leader, with two of its three products holding the top positions. Interestingly, prices per ounce are slightly lower for the 32-oz. sizes than for the 28-oz. and 44-oz. sizes.

4.2. Results

We now report our findings for the segment model when we apply it to the ketchup data. In Table 2, we report the parameter estimates for various specifications of heterogeneity ranging from one to three segments.[7] To control for potential heterogeneity in tastes across stores, we include a store-specific intercept in the conditional indirect utility of each of the four products. These parameters allow the size of the ketchup category to vary by store and should be

[6] We assign the factor price W_t the value from the corresponding month and then smooth the series: $W_t^s = 0.25W_{t-1} + 0.5W_t + 0.25W_{t+1}$.

[7] With the exception of the SUR results, all the models are over-identified. The 10 exogenous instruments are: the intercept, factor prices, feature ads, and displays (note that Heinz 44 is never displayed or featured). A pooled regression of prices on the instruments with brand intercepts yields an R-square of 0.45.

interpreted as differences with respect to store 9.[8] In the two- and three-segment models, the parameters representing store-specific mean utility of the inside goods are common across segments for reason of parsimony. Note that the GMM objective function values and degree of freedom reflect the joint estimation of the full model, which also includes the supply-side coefficients reported in Table 3. We separate the discussion of the demand and supply results to simplify the presentation. The first two columns contain the results for the seemingly unrelated regression (SUR) and the 3SLS regressions, respectively. Both approaches assume homogeneity (a single segment). As expected, we find that the SUR results, which do not instrument for the price endogeneity, yield a price-response parameter that is much lower than that of the 3SLS. This downward effect is consistent with the downward biases found in Berry et al. (1995), Besanko et al. (1998), Villas-Boas and Winer (1999), and Nevo (2001). Table 2 also implies that the price-response parameter estimated by 3SLS is less than (in absolute value) the segment-size-weighted average of the price-response parameters estimated in the multisegment models. This is consistent with the literature that discusses the consequences of ignoring unobserved heterogeneity (Chintagunta et al. 1991).

We now focus our attention on the remaining columns of Table 2, which report our findings for the multisegment models. The data identify three segments. Although not reported, our estimates for the four-segment model yield several insignificant parameters, including the probability of membership in the fourth segment. Thus, we conclude that the underlying consumer preferences generating our data are best represented by three segments and we will focus on the results of the three-segment model.

Comparing the brand-specific constants within each segment in Table 2, we note that segments 1 and 2 have stronger preference for the 28-oz. and 32-oz. sizes of Heinz, while segment 3 prefers Hunts 32. Segment 1 consumers account for roughly 33% of the shopping trips. This segment consists of price-insensitive shoppers who also do not respond to feature ads, but are somewhat responsive to displays.

[8] In the raw data we find that the no-purchase share varies on average across stores much more than the inside shares.

Table 2   **Parameter Estimates and Standard Errors for Alternative Models**

	Model													
	SUR		3SLS		2-segment				GMM 3-segment					
Attributes	Param	Se	Param	Se	Seg 1	Se 1	Seg 2	Se 2	Seg 1	Se 1	Seg 2	Se 2	Seg 3	Se 3
Price ($/oz.)	−77.58	1.63	−102.78	3.65	−57.46	0.16	−282.16	0.23	−44.58	2.20	−94.46	2.20	−214.14	4.68
Feature	0.46	0.05	0.34	0.06	0.60	0.05	1.38	0.09	0.15	0.16	1.16	0.16	0.02	0.36
Display	0.37	0.07	0.28	0.08	0.50	0.07	0.91	0.13	0.47	0.15	0.75	0.15	0.56	0.33
Heinz 32	0.69	0.08	1.80	0.17	0.09	0.05	7.18	0.06	−0.65	0.14	1.88	0.14	3.65	0.53
Hunts 32	0.08	0.08	1.17	0.16	−0.49	0.05	7.05	0.08	−0.67	0.15	−0.27	0.15	5.39	0.52
Heinz 28	1.08	0.10	2.38	0.20	0.40	0.05	6.68	0.05	0.05	0.19	1.92	0.19	1.64	0.42
Heinz 44	−0.54	0.09	0.67	0.18	−1.12	0.05	5.91	0.18	−1.09	0.21	−1.03	0.21	3.45	0.79
Prob	N/A	N/A	N/A	N/A	0.69	0.01	0.31		0.33	0.02	0.37	0.02	0.30	
					Estimate		Se		Estimate				Se	
Store intercept 1	0.71	0.05	0.71	0.05	0.91		0.06		0.91				0.07	
Store intercept 2	−0.03	0.05	−0.05	0.05	0.00		0.06		0.02				0.06	
Store intercept 3	0.36	0.05	0.35	0.05	0.48		0.06		0.48				0.06	
Store intercept 4	−0.61	0.06	−0.65	0.06	−0.65		0.07		−0.65				0.07	
Store intercept 5	0.19	0.06	0.18	0.06	0.28		0.07		0.29				0.07	
Store intercept 6	0.16	0.05	0.14	0.05	0.23		0.06		0.24				0.06	
Store intercept 7	−0.50	0.06	−0.52	0.07	−0.58		0.08		−0.59				0.08	
Store intercept 8	0.37	0.05	0.34	0.05	0.47		0.06		0.49				0.06	
GMM objective (df)	0.6720 (0)		1.09 (58)		2.33E-4 (50)				2.19E-4 (42)					

Segment 2, accounting for about 37% of the shopping trips, is moderately price sensitive, but highly responsive to feature ads and displays. This segment seems to fit the profile of the "time-starved" consumer: a consumer who does not pay that much attention to relative prices when choosing among brands, but who can be persuaded to buy one brand over the others when that brand is promoted through feature advertising or displays. Finally, segment 3, accounting for about 30% of the trips, has very high price sensitivity and moderate sensitivity to displays. Our analysis of shares in the three segments reveals that segment 3 has the lowest incidence of ketchup purchases on average, and segment 1 has the highest.

Several of the estimated store-specific intercepts are significantly different from zero, indicating differences in the average size of the ketchup category across stores. The estimated store 1 parameter is largest, and store 4 parameter is smallest. In the subsequent section, we use store 1 as an example for illustrating different price discrimination strategies.

In Table 3 we present the marginal cost estimates for the three-segment model. We find that the

coefficients on the prices of ingredients, containers, and labor have the expected positive sign.

We now take our three-segment model and compute price elasticities, which we report in Table 4 as means over all store-weeks. All of the own elasticities are greater than one in magnitude, which is consistent with static oligopoly behavior. Hunts 32 has the largest price elasticity of demand, while Heinz 44 has the smallest elasticity. Heinz 44 tends to be the least price responsive to changes in competitors' prices. In terms of stimulating consumers to switch away from the no-purchase alternative, change in the price of

Table 3   **Marginal Cost Parameter Estimates for the Three-Segment Model**

Factors	Param	Se
Heinz 32	−0.373	0.021
Hunts 32	−0.356	0.021
Heinz 28	−0.374	0.021
Heinz 44	−0.371	0.021
Labor	0.005	0.001
Container (*10E-3)	0.384	0.074
Ingredients (*10E-3)	30.466	1.670

Table 4   Mean Own-Price Elasticities for the Three-Segment Model

| Change in price | Change in share | | | | |
	Heinz 32	Hunts 32	Heinz 28	Heinz 44	No purchase
Heinz 32	−3.14	0.09	0.22	0.04	0.23
Hunts 32	0.17	−3.76	0.17	0.05	0.17
Heinz 28	0.27	0.10	−3.12	0.05	0.15
Heinz 44	0.17	0.13	0.21	−2.54	0.03

Heinz 32 has the greatest impact, changes in Hunts 32 and Heinz 28 have a moderate impact on category demand, and Heinz 44 has almost no impact.

In Table 5 we provide segment-specific elasticities, which give additional insight into the substitution patterns reported above. Because the substitution patterns within a segment exhibit the IIA property, we do not report the cross-elasticities. Looking at Table 5, we see that the magnitudes of the own-price elasticities differ substantially across segments, reflecting the large differences in price response. Segment 3 is generally the most price elastic, and segment 1 the least price elastic. This suggests that if the retailer or the manufacturer could engineer segment-specific prices, there might be significant gains from directing targeted price cuts (perhaps in the form of coupons) to consumers in segment 3. Of the four products, Heinz 32 and Hunts 32 have the lowest price elasticities across segments.

As a final point about Table 5, notice that the rankings of the own-price elasticities by segment are the same for each brand. That is, for each brand, the price elasticity of demand is largest for segment 3 and smallest for segment 1. This symmetry in the elasticity rankings across the segments is significant when we explore the economics of price discrimination in the next section.

Table 5   Mean Segment Own-Price Elasticities for the Three-Segment Model

| | Own-price elasticities | | | |
	Overall	Seg 1	Seg 2	Seg 3
Heinz 32	−3.14	−1.85	−3.73	−9.29
Hunts 32	−3.76	−1.75	−4.01	−8.79
Heinz 28	−3.12	−2.18	−4.87	−11.39
Heinz 44	−2.54	−2.03	−4.41	−10.00

5.   In-Store Strategies to Capture Consumer Surplus

The economic model estimated in the paper assumes that the retailer sets a uniform profit-maximizing price, given the wholesale prices. Thus, the retailer and the manufacturers are assumed to know the underlying segment structure of demand, but they are unable to price discriminate. In this section, we use our empirical estimates of ketchup demand to study a form of third-degree price discrimination. Third-degree price discrimination is the practice of charging segment-specific prices for consumer segments that differ according to their price elasticities of demand.

We study two forms of third-degree price discrimination. First, we analyze third-degree price discrimination engineered by the retailer. In this analysis, the retailer is assumed to set segment-specific prices to maximize retail profits. Second, we study third-degree price discrimination engineered by manufacturers. In this analysis, manufacturers effectively set segment-specific wholesale prices by offering scan-back discounts that induce the retailer to charge targeted retail prices. An issue that arises in this analysis is whether the decision to engage in price discrimination results in a prisoner's dilemma, in which all manufacturers offer targeted prices without increasing their profitability.

Our goal is to identify potential increases in profitability if retailers and/or manufacturers used their knowledge of consumer segments to implement price discrimination strategies. Another question of interest is the extent to which retailer and manufacturer incentives are aligned when one party in the channel engineers the price discrimination scheme. Note that because we assume that customer-specific histories are not available, the targeting scheme must rely on the current purchase decision of a customer. We compare the profit implications of this checkout-counter couponing scheme to one in which retailers and/or manufacturers are able to classify consumers perfectly into segments. Note that we assume all coupons are redeemed. In this respect, the analysis is intended to provide an upper bound on the potential gains from targeted coupons.

Table 6 Equilibrium Prices and Expected Per-Trip Profits Under Uniform Pricing

Product	Wholesale prices ($/oz.)	Retail prices ($/oz.)	Retail profits ($/oz.)	Manufacturer profits ($/oz.)
Heinz 32	0.029	0.048	0.049	0.058
Hunts 32	0.029	0.041	0.029	0.008
Heinz 28	0.031	0.052	0.050	0.055
Heinz 44	0.023	0.045	0.026	0.014

5.1. Uniform Pricing Equilibrium

To provide a benchmark for the subsequent counterfactual analyses, we use the parameter estimates of the three-segment model for the ketchup data to compute equilibrium prices under the assumption that all consumers, regardless of segment, face a uniform price for each brand. Table 6 shows these prices, along with the corresponding profits for the retailer and the two manufacturers (Heinz and Hunts). In this table, prices are expressed in dollars per ounce, while profits are in dollars per ounce per store trip. Because the estimated model assumes uniform pricing, it is not surprising that the optimal prices are close to the actual average prices shown in Table 1.

To assess the quality of targeting information available to the retailer and/or manufacturers, we use an empirical analog of Bayes rule to compute the posterior probability of segment membership conditional on current purchases. Thus, each of the five possible choices (four brands and no purchase) induces a conditional demand system to which firms may now target different prices. The more the posterior probabilities are closer to zero or one, the more informative is the current purchase for targeting. Across the nine stores in our data we find substantial differences in the informativeness of the current purchase, hence in the profit gains from targeting. For the subsequent coupon analysis, we only report results for consumers in store 1.

5.2. Third-Degree Price Discrimination by the Retailer

As a benchmark, we begin by considering the case in which the retailer can perfectly distinguish among the three segments identified in our empirical analysis and can charge a different price for each segment, a scenario we refer to as *perfect classification*. Table 7

shows the equilibrium prices and profits that arise in this scenario.[9] In light of the segment-level price elasticities shown in Table 5, it is not surprising that the retailer lowers the prices charged to segment 3 and increases the prices to segment 1 as compared to uniform pricing. Responding to the scheme of discriminatory retail margins chosen by the retailer, Heinz lowers the wholesale prices on its brands, while Hunts raises its wholesale price. Comparing the retailer's profits under uniform pricing (as shown in Table 6) versus price discrimination with perfect classification (Table 7), we see that retail profits rise on all Heinz brands but fall on the Hunts product. Overall retail profits increase by 21%.

In practice, perfect classification of consumers to segments could not be achieved unless the retailer maintained purchase histories for individual consumers. Rossi et al. (1996) show that the ability to identify a consumer's preferences increases with the amount of past purchase information available and that even with fairly limited data, one may be able to predict these preferences quite accurately. However, without a consumer microdatabase such as loyalty cards, the only information typically available to determine a consumer's segment membership is the current transaction at the checkout counter.[10] Accordingly, in this section we consider third-degree price discrimination based on *imperfect classification*. Under this mechanism, the retailer observes the consumer's current purchase within the category and based on that purchase, it generates (via Bayes Rule) posterior segment-membership probabilities for that consumer.[11] These posteriors are, in turn, used as a basis for issuing customized Catalina-like coupons to the consumer at the checkout counter, redeemable on a future purchase occasion. Coupons would be offered for all products in the category, and their face values would be conditioned on the particular choice made

---

[9] The prices shown in Table 7 take into account equilibrium wholesale price responses of manufacturers to the new retail prices implied by the retailer's price discrimination strategy.

[10] And, of course, it is the lack of such microlevel data that provides a rationale for estimating segment-level demands based on aggregate data, as we have done in this paper.

[11] Kamakura and Russell (1989) discuss this logic for a model estimated with household panel data.

Table 7   **Equilibrium Prices and Expected Per-Trip Profits with Retail Price Discrimination: Perfect Classification**

Brand	Wholesale prices ($/oz.)	Shelf prices ($/oz.)	Prices net of coupon ($/oz.)			Retail profits ($/oz.)	Manufacturer profits ($/oz.)
			Seg 1	Seg 2	Seg 3		
Heinz 32	0.027	0.068	0.068	0.046	0.028	0.065	0.068
Hunts 32	0.033	0.077	0.077	0.055	0.042	0.024	0.012
Heinz 28	0.031	0.066	0.066	0.045	0.026	0.056	0.058
Heinz 44	0.022	0.070	0.070	0.048	0.030	0.041	0.033

by the consumer in the current transaction (including the choice not to purchase in the category at all). The classification scheme is imperfect because a single transaction can, at best, only allow the retailer to update the probability that a consumer belongs to particular segments, rather than perfectly identify the consumer's segment membership.

Table 8 shows the equilibrium prices and profits that arise when the retailer engineers a scheme of third-degree price discrimination based on imperfect classification using checkout-counter information. Effectively, this procedure generates five "purchase segments" based on the observed current purchase—one for each of the alternatives purchased, including no purchase. Note that there is a separate set of equilibrium prices for each segment. The term "shelf price" in this table is the highest segment-specific price for each of the four products. Based on the observed current purchase, a customized coupon is offered to the consumer, which effectively reduces the price of each product (on the next purchase occasion) to the purchase-segment specific equilibrium price level.[12] Interestingly, as compared to uniform pricing, we observe that prices for all products generally rise for consumers who currently purchase Hunts 32, Heinz 28, or Heinz 44. This occurs because consumers who purchase these products have a higher likelihood of coming from segment 1, the least price-sensitive segment. Similarly, prices are lowest for consumers who choose no purchase because these consumers are most likely to be from segment 3, the most

price-sensitive segment. Responding to the scheme of discriminatory retail margins chosen by the retailer, Heinz lowers the wholesale price on one of its products, maintains the price on one product, and raises the price on the third product, whereas Hunts lowers its price.

Comparing Tables 6 and 8, we see that retail profits rise substantially for Hunts 32 and marginally for the Heinz product line. Overall, with imperfect classification, the retailer's profits increase by 11% relative to the uniform pricing equilibrium. Thus, even when a scheme of discriminatory prices is based on a single transaction at the checkout counter, profits go up as compared to uniform pricing. Indeed, checkout-counter information can generate 53% of the gains that are attainable under perfect classification, without the high costs of collecting and maintaining historical information. The general message here is that the use of even small amounts of information gleaned from segment-level demand estimates based on aggregate data can potentially unlock nontrivial sources of additional profitability for a retailer. We also computed a constrained solution in which a coupon on only one product is issued at a time, and find that scheme produces incremental profits relative to uniform pricing.[13]

Interestingly, retailer-engineered price discrimination does not unambiguously benefit the manufacturers. Both manufacturers are better off under perfect classification, but worse off under imperfect

[12] As under perfect classification, the prices shown in Table 8 take into account equilibrium wholesale price responses of manufacturers to the new retail prices implied by the retailer's price discrimination strategy.

[13] In practice, one might expect retailers to issue only a single coupon on one of the products in the category, instead of separate coupons for each product. We simulated this scenario and found that the retailer could still generate a 5% profit gain if a coupon were issued for Heinz 32.

Table 8 **Equilibrium Prices and Expected Per-Trip Profits with Retail Price Discrimination: Imperfect Classification**

Brand	Wholesale prices ($/oz.)	Shelf prices ($/oz.)	Prices net of coupon for purchase segment ($/oz.)					Retail profits ($/oz.)	Manufacturer profits ($/oz.)
			Heinz 32	Hunts 32	Heinz 28	Heinz 44	No purchase		
Heinz 32	0.029	0.056	0.047	0.054	0.050	0.056	0.047	0.053	0.058
Hunts 32	0.026	0.057	0.052	0.046	0.055	0.057	0.036	0.040	0.004
Heinz 28	0.029	0.059	0.050	0.058	0.054	0.059	0.050	0.054	0.055
Heinz 44	0.024	0.055	0.050	0.053	0.053	0.055	0.047	0.024	0.013

classification. We also see that total channel profits rise with perfect classification, and the retailer continues to get a little more than half of the total channel profits. As expected, total channel profits under imperfect classification are lower than under perfect classification.

5.3. Third-Degree Price Discrimination by the Manufacturers

Now we turn to the case in which manufacturers engineer price discrimination by using knowledge of segment-specific demands. To implement third-degree price discrimination, manufacturers set wholesale prices contingent on the consumer's purchase segment. Conditional on these targeted wholesale prices, the retailer determines targeted retail prices to charge to each segment. To focus on the case of price discrimination engineered by manufacturers, we assume the retailer will not price discriminate if manufacturers choose not to offer targeted wholesale prices. In practice, one could think of this pricing scheme as a form of "targeted scanbacks." The manufacturer issues a contract with per-unit discounts on the wholesale price, contingent on the consumer purchase segments and prespecified targeted retail prices for each segment. By setting the prespecified retail prices to the retailer's profit-maximizing levels implied by the Stackelberg equilibrium, the retailer will accept the terms. The implementation of prespecified targeted retail prices occurs via checkout-counter coupons, much like the process assumed in the case of retailer-engineered price discrimination discussed previously.

When price discrimination is engineered by manufacturers, the pricing game becomes more complicated because the benefits of price discrimination for a manufacturer depend on whether the other manufacturer price discriminates as well. When a monopolist price discriminates, its profits must necessarily go up relative to the uniform pricing case, but as Corts (1998) shows, price discrimination by oligopoly firms can sometimes lead to lower profits. The intuition is that if a firm can price discriminate, it will be tempted to use discriminatory price cuts to steal demand from rivals in its "weak" market segments—market segments in which demand for the firm's product is highly price elastic—without compromising profit margins in its "strong" market segments—market segments in which demand for the firm's product is relatively less price elastic. The ability to discriminate lowers the "cost" of price cutting, causing the firm to become a more aggressive price competitor in its "weak" segments. When the "weak" segments for one firm are the "strong" segments of the rival firm (and vice versa), then price discrimination in an oligopoly can lead to what Corts (1998) calls "all-out-competition": Relative to the uniform pricing case, prices go down in all market segments. In such cases, oligopolists can often face a prisoner's dilemma: Each firm has a dominant strategy to adopt price discrimination, but adoption of price discrimination by all firms results in a reduction of industry profit relative to the no-discrimination case.

Whether the conditions under which price discrimination leads to "all-out-competition" hold in real consumer products categories is a question of significant practical importance. The payoff matrix shown in Figure 1, which is based on the equilibrium prices shown in Table 9, suggests that third-degree price discrimination by manufacturers in the ketchup category does *not* lead to "all-out-competition" when consumers can be perfectly classified into segments. Figure 1 shows that price discrimination by both firms

Figure 1   Expected Manufacturer Profits Under Manufacturer-Engineered Price Discrimination: Perfect Classification

		Heinz	
		No Price Discrimination	*Price Discrimination*
Hunts	No Price Discrimination	Hunts: $0.0081 Heinz: $0.1266	Hunts: $0.0079 Heinz: $0.1320
	Price Discrimination	Hunts: $0.0080 Heinz: $0.1353	Hunts: $0.0137 Heinz: $0.1773

is the equilibrium outcome, and that this price discrimination leads to increases in profit by both firms. This result is related to the nature of market segmentation revealed by our demand estimation. As just discussed, the case in which price discrimination would be expected to induce all-out competition is one in which there is an asymmetry in firms' "weak" and "strong" markets. However, the market segmentation in ketchup, as illustrated by Table 5, reveals that both Heinz and Hunts would rank the same group of consumers (segment 1) as "strong" and the same group as "weak" (segment 3). It is in this case that price discrimination by rival manufacturers would be expected to increase industry profits.

With imperfect classification, consumers' current purchase is used to assign them to "purchase segments." In this case, manufacturer-engineered price discrimination leads to higher profits for both Hunts and Heinz, as shown in Figure 2. Interestingly, coupons permit Hunts to increase its profits by 37% with imperfect classification and 69% with perfect classification, suggesting that checkout-counter information enables Hunts to achieve over half of the total potential gains from couponing. Gains for Heinz

are more modest—40% with perfect classification and 4% with imperfect classification.

Comparing manufacturer-engineered discriminatory prices, shown in Table 8, to retailer-engineered discriminatory prices (for imperfect classification), shown in Table 10, we see that the segment-specific retail prices for Heinz 32 and Heinz 28 are quite similar. By contrast, the prices of Hunts 32 and Heinz 44 are generally higher when manufacturers design the pricing scheme. This reflects the difference between the retailer's incentives (category profit maximization) and the manufacturers' incentives.

We also see that while total channel profits are about the same for the two cases, the joint share of the manufacturers is higher when they engineer the price discrimination rather than when the retailer does (for example, 55% versus 48% with perfect classification).

5.4. Comparison to the Previous Literature
Our results suggest that manufacturer-engineered targeted coupons can increase the profitability of the manufacturers and the retailers even with imperfect targetability. This conclusion differs sharply from the extant analytical findings that competitive targeted

Figure 2   Expected Manufacturer Profits Under Manufacturer-Engineered Price Discrimination: Imperfect Classification

		Heinz	
		No Price Discrimination	*Price Discrimination*
Hunts	No Price Discrimination	Hunts: $0.0081 Heinz: $0.1266	Hunts: $0.0048 Heinz: $0.1432
	Price Discrimination	Hunts: $0.0119 Heinz: $0.1307	Hunts: $0.0111 Heinz: $0.1316

couponing leads to a prisoner's dilemma, reducing all manufacturers' profits (Shaffer and Zhang 1995), and that the prisoner's dilemma ensues as the precision of targetability increases (Chen et al. 2001). The main distinctions between our results and those of analytical models are our treatments of product differentiation and the econometric error. We also allow for the possibility of category expansion, although we find evidence of gains irrespective of this aspect.

Shaffer and Zhang (1995) use a model of horizontal differentiation à la Hotelling and base targeting decisions on the assumption that one can use choice data to learn about consumer locations along the unit interval. The prisoner's dilemma arises because, using the intuition of Corts (1998), their model leads to all-out competition. In the Hotelling model, one firm's strong segment (consumers that are physically close to it along the interval and hence have high willingness to pay) constitutes the competitor's weak segment. This asymmetry entices firms to target one another's strong segments with low prices.

In contrast, our model allows for both horizontal and vertical differentiation. In Equation (1), we write the conditional indirect utility in our model obtained by some consumer h in segment k who consumes product j. We can rescale this expression by dividing by α^k and rewriting the rescaled idiosyncratic error as $\varepsilon_{hj}/\alpha^k = \tau^k(z^h - z^j)^2$. This gives us

$$\widetilde{u_{hj}^k} = X_j t^k - p_j + \tau^k(z^h - z^j)^2.$$

Using a result in Anderson et al. (1992), this formulation is completely analogous to the spatial model used by Shaffer and Zhang (1995).[14] In this setting, $X_j t^k$ captures vertical product differentiation and the econometric error term $\tau^k(z^h - z^j)^2$ captures horizontal product differentiation, where z^h measures a consumer's ideal point along some latent attribute dimension and z^j measures the location of product j

in that same dimension. As in the empirical literature (Rossi et al. 1996), we base targeting decisions on the ability to learn about t^k, consumers' heterogeneous tastes for vertical attributes. The vertical differentiation allows consumers to have high willingness to pay for products that are economically "distant" along the horizontal dimension, which may counter the all-out-competition result. Data-based methods do not enable one to learn about the econometric error term, $\tau^k(z^h - z^j)^2$, empirically ruling out targeting along the horizontal dimension.

5.5. Discussion and Extensions

So far, we compare targeting in the case of perfect classification (synonymous with having a vast consumer panel) to the case of a single checkout-counter purchase, where the latter reflects the lack of availability of panel data. While expected gains are lower in the latter case, we still find benefits from couponing. These benefits accrue from two sources. First, targeting increases efficiency as prices and products can be realigned more carefully with consumer taste segments. Second, targeting can increase the category size by offering lower prices to consumers with a high reservation value.

A potential limitation of our aggregate data-based targeting strategy is the lack of consumer inventory information. In particular, the gains due to category expansion may be overly optimistic if nonpurchase also reflects inventories, which could change consumer willingness to pay over time. Omission of the inventory effect may overstate the predicted demand response to couponing. Although not reported, our analysis of household panel data for ketchup, taking into account the effects of inventory, shows small changes in price elasticities, and therefore in targeted prices, across weeks with high versus low inventories. These results suggest that either the economic impact of inventories is small, or that it is not well identified even when panel data are available.

Because we cannot control for inventories, we must consider how ignoring them could affect our targeting strategy. In general, we expect inventories to be problematic in categories with long purchase cycles. At the beginning of the cycle, a consumer receives coupons based on her purchase. However, during subsequent trips before the end of the cycle,

---

[14] It can be shown that the density of consumer locations must have the form

$$g(z) = N(4b\tau)^{n-1}(n-1)! \frac{\Pi_{i=1}^{n-1} \exp(-4b\tau z^i)}{[1 + \sum_{j=1}^{n-1} \exp(-4b\tau z^i)]^n}$$

to make the comparison between the logit and the spatial model. In this setting, N is the size of the market, b is a positive constant denoting proximity of the variants, and n is the number of alternatives.

Table 9 Equilibrium Prices and Expected Per-Trip Profits with Manufacturer Price Discrimination: Perfect Classification

	Wholesale prices ($/oz.)			Shelf prices ($/oz.)	Prices net of coupon ($/oz.)			Retail profits ($/oz.)	Manufacturer profits ($/oz.)
Brand	Seg 1	Seg 2	Seg 3		Seg 1	Seg 2	Seg 3		
Heinz 32	0.041	0.031	0.018	0.069	0.069	0.047	0.026	0.059	0.077
Hunts 32	0.049	0.036	0.030	0.077	0.077	0.052	0.039	0.018	0.014
Heinz 28	0.040	0.030	0.017	0.068	0.068	0.046	0.025	0.055	0.071
Heinz 44	0.043	0.033	0.020	0.071	0.071	0.049	0.028	0.024	0.029

she would reveal herself as a nonpurchaser, potentially entitling her to better coupons and offsetting the potential benefits of couponing. To alleviate this problem, we use the channel model to compute a restricted couponing scheme whereby retailers issue optimal coupons subject to the constraint that the nonpurchasers do not receive one (e.g., pay the shelf price). Under this scenario, consumers only receive coupons based on purchases, so that we avoid the confound between inventory effects versus low reservation value for ketchup. We find that this couponing scheme improves the retailer's expected profits by 4%. These expected gains are a lower bound on the potential for couponing using aggregate data because we eliminate the potential for gains from category expansion.

6. Conclusions and Limitations

In this paper, we explore the extent to which estimated taste heterogeneity offers firms an opportunity to price discriminate profitably by charging different prices across consumer segments. The context of our study is a packaged-goods market in which two large competing manufacturers sell through a large supermarket retailer. Using aggregate retail data, we estimate a discrete choice demand system with latent consumer segments, jointly with a structural channel model of uniform pricing by manufacturers and the retailer. The use of aggregate data, in contrast with use of household panel data, is a novel approach relative to the literature, and is consistent with the kinds of data to which retailers typically have access. However, it implies that firms can classify consumers into segments only imperfectly. Note that if one had access to store-specific household panel data, one could carry out an analogous competitive couponing exercise as described herein.

We then analyze price discrimination strategies as counterfactual experiments, using the estimated demand and supply-side parameters. We explore both manufacturer-engineered and retailer-engineered segment pricing strategies, using the channel model to control for competitive reactions. We find that in equilibrium it is profitable for the retailer to price discriminate by issuing segment-specific discount coupons to consumers. Despite the imperfect classification of consumers to segments, the potential gains from price discrimination are substantial. However, at the new equilibrium manufacturers are not unambiguously better off as compared with the uniform pricing equilibrium. When manufacturers are allowed to engineer price discrimination and segment classification is imperfect, we find that issuing coupons

Table 10 Equilibrium Prices and Expected Per-Trip Profits with Manufacturer Price Discrimination: Imperfect Classification

	Wholesale prices ($/oz.) for purchase segment					Shelf prices ($/oz.)	Prices net of coupon ($/oz.) for purchase segment					Retail profits ($/oz.)	Manufacturer profits ($/oz.)
Brand	Heinz 32	Hunts 32	Heinz 28	Heinz 44	No purchase		Heinz 32	Hunts 32	Heinz 28	Heinz 44	No purchase		
Heinz 32	0.031	0.025	0.031	0.035	0.029	0.061	0.049	0.049	0.051	0.061	0.047	0.053	0.060
Hunts 32	0.045	0.025	0.047	0.048	0.030	0.077	0.072	0.044	0.076	0.077	0.041	0.025	0.011
Heinz 28	0.031	0.029	0.032	0.037	0.030	0.065	0.050	0.058	0.056	0.065	0.051	0.052	0.057
Heinz 44	0.040	0.017	0.042	0.043	0.027	0.072	0.068	0.043	0.071	0.072	0.051	0.019	0.013

based on purchases is a dominant strategy for both manufacturers. We also find that manufacturers' share of total channel profits is higher when they initiate the price discrimination.

There are a number of possible extensions of the techniques that we present here. For instance, in our analysis, we have assumed that all retail stores face local markets with an identical segment structure. It is straightforward to adapt the model to the possibility that retail stores face different segment structures or different relative segment sizes. Future research could explore the potential of market-level demographic information to explain differences in segment sizes across stores.

Our use of a finite mixture representation of heterogeneity was guided by considerations of computational tractability and managerial interpretability of the estimated latent segments. Continuous random-effects models of heterogeneity (Allenby and Rossi 1999) have been found to provide superior fit to the data, especially at the tails of the heterogeneity distribution. Future work could assess the implications of employing alternative representations of heterogeneity for the profitability of price-discrimination strategies.

The inherently lower information content of aggregate data also introduces certain limitations. We discuss the lack of consumer inventory information, which could lead to a time-varying elasticity of demand if consumers' willingness to pay varies with their stock of ketchup. We propose a restricted couponing policy when one does not have information on inventories. However, conditioning on inventory information, when it matters, would likely improve expected profits from targeting. Future research could explore how to time and design targeted coupons when consumer elasticities vary substantially across store trips. Another concern with the use of aggregate data is the inability to incorporate multiple unit purchases. This concern is not severe in our empirical application to the ketchup category, wherein almost all purchases are of single units.

Acknowledgments

Appendix. Model Performance with Simulated Data

We demonstrate via numerical simulations that the econometric model described in §§2 and 3 can recover the structure of the underlying segments from realistic aggregate store-level data. Our simulated market consists of five segments. Within each segment, consumers are identical and make logit choices across six brands and a no-purchase alternative. The impact of all fixed product attributes and brand reputations is captured by brand-specific constants, which differ across segments. Prices affect utilities and hence choices, and are assumed to be exogenous. (Because the objective here is to show recoverability of demand heterogeneity, we do not deal with issues of price endogeneity.) Prices of the six brands are generated from independent uniform distributions. The price parameter varies across segments. We assume that data are pooled across 50 stores and 100 weeks. These appear to be reasonable choices for a medium- to large-sized supermarket chain. We generate store trips exogenously at the individual level, and the choice parameters and prices result in discrete choices from the set of six products and no purchase. These are aggregated to obtain weekly brand shares.

The baseline demand model described in §§2 and 3 is estimated using GMM. In evaluating the quality of model estimates, we focus on two metrics that are relevant for the pricing application: price elasticities and estimated equilibrium margins. These are considered both for the aggregate market and separately for each segment. We use mean absolute percentage deviation (MAPD) as the measure of bias. It is relevant to note that this measure includes both bias and variance, because only one sample was used for the estimation (due to the computational cost of multiple replicates). A summary of the findings is presented here; detailed tables are available from the authors.

The MAPD (across six brands) in aggregate own elasticities is 2.7%, while the MAPD in aggregate cross elasticities is 12.4%. Thus, the aggregate elasticities are recovered quite well by the model. Turning to the segment-level performance, we note that cross-price elasticities are restricted within segments because of the IIA property of the logit model. With the exception of segment 1, the smallest segment, the own-price elasticities are estimated with MAPD ranging from 5.4% to 13.4%. While the MAPD in cross elasticities is larger (ranging from 19.9% to 46.9%), the absolute values of the cross elasticities to which these apply are quite small. We also compare true margins with estimated equilibrium margins by brand for each segment and for the aggregate market. Results indicate that the estimated margins are very close to the true margins in aggregate (absolute percentage deviation ranges from 0% to 12% across brands) and for each segment, with the exception of segment 1.

References

Ackerberg, D. 2001. Empirically distinguishing informative and prestige effects of advertising. *RAND J. Econom.* **32** 316–333.

Anderson, S., A. de Palma, J. F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* MIT Press, Cambridge, MA.

Allenby, G. M., P. E. Rossi. 1999. Marketing models of consumer heterogeneity. *J. Econometrics* **89** 57–78.

Berry, S. 1994. Estimating discrete-choice models of product differentiation. *RAND J. Econom.* **25** 242–262.

——, M. Carnall, P. T. Spiller. 1997. Airline hubs: Costs, markups and the implications of customer heterogeneity. Working paper, Yale University, New Haven, CT.

——, J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63** 841–890.

Berto Villas-Boas, S. 2002. Vertical contracts between manufacturers and retailers: An empirical analysis. Working paper, University of California, Berkeley, CA.

Besanko, D., J. P. Dubé, S. Gupta. 2003. Retail pass-through on competing brands. *Marketing Sci.* Forthcoming.

——, S. Gupta, D. Jain. 1998. Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Sci.* **44** (11) 1533–1547.

Bucklin, R. E., S. Gupta. 1999. Commercial use of UPC scanner data: Industry and academic perspectives. *Marketing Sci.* **18**(3) 247–273.

Chen, Y., C. Narasimhan, Z. J. Zhang. 2001. Individual marketing with imperfect targetability. *Marketing Sci.* **20** 23–43.

Chintagunta, P. 1993. Investigating purchase incidence, brand choice, and purchase quantity decisions of households. *Marketing Sci.* **12**(2) 184–209.

——. 2002. Endogeneity and heterogeneity in a probit demand model: Estimation using aggregate data. *Marketing Sci.* **20** 442–456.

——, D. Jain, N. J. Vilcassim. 1991. Investigating heterogeneity in brand preferences in logit models for panel data. *J. Marketing Res.* **28** 417–428.

Choi, S. Chan. 1991. Price competition in a channel structure with a common retailer. *Marketing Sci.* **10** 271–296.

Corts, K. 1998. Third degree price discrimination in oligopoly: All-out competition and strategic commitment. *RAND J. Econom.* **29**(Summer) 306–323.

Dreze, X., S. J. Hoch, M. E. Purk. 1994. Shelf management and space elasticity. *J. Retailing* **70** 301–326.

Gupta, S., P. K. Chintagunta, A. Kaul, D. R. Wittink. 1996. Do household scanner data provide representative inferences from brand choices: A comparison with store data. *J. Marketing Res.* **33**(4) 383–398.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* **50** 1029–1054.

Kamakura, W. A., G. J. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* **26** 379–390.

McFadden, D. 1981. Econometric models of probabilistic choice. C. Manski, D. McFadden. *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge, MA.

Moorthy, S. 1984. Market segmentation, self selection, and product line design. *Marketing Sci.* **3**(4) 288–307.

Narasimhan, C. 1984. A price discrimination theory of coupons. *Marketing Sci.* **3** 127–147.

Nevo, A. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* **69** 307–340.

Rossi, P. E., R. E. McCulloch, G. M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Sci.* **15** 321–340.

Shaffer, G., Z. J. Zhang. 1995. Competitive coupon targeting. *Marketing Sci.* **14** 395–416.

Slade, M. E. 1995. Product rivalry with multiple strategic weapons: An analysis of price and advertising competition. *J. Econom. Management Strategy* **4**(3) 445–476.

Sudhir, K. 2001. Structural analysis of competitive pricing in the presence of a strategic retailer. *Marketing Sci.* **20** 244–264.

Varian, H. 1989. Price discrimination. R. Schmalensee, R. Willig, eds. *Handbook of Industrial Organization.* North-Holland, Amsterdam, The Netherlands.

Villas-Boas, M., R. Winer. 1999. Endogeneity in brand choice models. *Management Sci.* **45**(10) 1324–1338.

Wedel, M., W. Kamakura. 1998. *Market Segmentation: Conceptual and Methodological Foundations*. Kluwer Academic Publishers, Boston, MA.

Wheat, R. D., D. G. Morrison. 1990. Assessing purchase timing models: Whether or not is preferable to when. *Marketing Sci.* **9**(2) 162–170.

Zenor, M. J., R. Srivastava. 1993. Inferring market structure with aggregate data: A latent segment logit approach. *J. Marketing Res.* **30** 369–379.