



Marketing Science

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Accounting for Primary and Secondary Demand Effects with Aggregate Data

Harikesh Nair, Jean-Pierre Dubé, Pradeep Chintagunta,

To cite this article:

Harikesh Nair, Jean-Pierre Dubé, Pradeep Chintagunta, (2005) Accounting for Primary and Secondary Demand Effects with Aggregate Data. Marketing Science 24(3):444-460. <https://doi.org/10.1287/mksc.1040.0101>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

© 2005 INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Accounting for Primary and Secondary Demand Effects with Aggregate Data

Harikesh Nair, Jean-Pierre Dubé, Pradeep Chintagunta

Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, Illinois 60637
{hnair@gsb.uchicago.edu, jdube@gsb.uchicago.edu, pradeep.chintagunta@gsb.uchicago.edu}

Discrete choice models of aggregate demand, such as the random coefficients logit, can handle large differentiated products categories parsimoniously while still providing flexible substitution patterns. However, the discrete choice assumption may not be appropriate for many categories in which we expect consumers may purchase more than one unit of the selected item. We derive the aggregate demand system corresponding to a discrete/continuous household-level model of demand. We also propose a method-of-simulated-moments procedure that provides consistent estimates of the structural parameters when only aggregate data are available. The procedure also enables the researcher to control both for the potential endogeneity of marketing variables as well as potential heterogeneity in consumer tastes. Using our aggregate estimates, we can measure the decomposition of price elasticities into incidence, brand choice, and purchase quantity components. We also propose several empirical tests to assess the validity of the discrete/continuous demand system versus that of the logit model. In several simulation experiments, we demonstrate the robustness of this model across datasets in which quantity choices may or may not be important. Our empirical calibration to store-level data in the refrigerated orange juice category indicates a considerable improvement in fit of the observed aggregate sales using the discrete/continuous model.

Key words: discrete/continuous demand system; logit demand system; aggregate data; price endogeneity; primary and secondary demand

History: This paper was received November 3, 2003, and was with the authors 2 months for 2 revisions; processed by Tulin Erdem.

1. Introduction and Background

Recent work in marketing with aggregate store and chain level data has increasingly resorted to the logit demand system (Besanko et al. 1998, Sudhir 2001). Unlike traditional log-log or AIDS models, the logit is parsimonious and is thus attractive for categories with a large number of differentiated products. However, the logit formulation (and other discrete choice models or DCMs) makes the implicit assumption of single-unit purchases by consumers and could suffer from a misspecification problem if multiple-unit purchases are frequent. To understand the source of misspecification one must examine the derivation of the aggregate demand system. Since consumers are constrained to purchase a single unit of their chosen alternative, the market share of a given product is obtained by integrating the heterogeneity distribution over the mass of consumers who are expected to choose that product. Parameter estimation then proceeds by matching observed market shares to the model-predicted market shares (c.f., Berry et al. 1995). If individual consumers purchase multiple units, however, simply integrating the product-choice probabilities over the distribution of heterogeneity will not

give the correct product market-shares. Rather, one must integrate the brand *and* quantity choices over the distribution of heterogeneity.¹ Ignoring the quantity choices leads to a misspecified demand system, which could generate incorrect sales forecasts. From a more managerial perspective, the model may also generate incorrect estimates of the brand-share elasticities, which would generate misleading recommendations for category pricing and margin decisions.²

When household data are available, the literature has proposed various models that accommodate purchase quantities within the context of the DCM (e.g., Krishnamurthi and Raj 1988, Chiang 1991,

¹ An exception would be the case in which consumers' quantity choices are independent of brand-choice decisions, and the quantity decision does not depend on prices and other marketing variables at the time of purchase. Both are unlikely to be the case, and are strong assumptions to impose a priori.

² If households are actually stocking up when prices are low, then ignoring this can result in overstating the long-run effect of prices on sales, leading to overstated own price elasticities. This requires a dynamic model of consumer demand (Erdem et al. 2003), for which individual-level data are better suited. We ignore this issue at least for the purposes of this paper.

Chintagunta 1993, Dillon and Gupta 1996, Arora et al. 1998, Mehta et al. 1998).³ While individual-level data are preferred, in many instances such disaggregate information may not be available. The wider availability of aggregate data, at the store, chain, or market-level necessitates the development of corresponding aggregate models. We therefore derive the corresponding aggregate demand system to the individual-level choice and quantity models of the type used by Chiang (1991) and Chintagunta (1993). These models approximate the quantity choice by assuming products are divisible; although the approximation has been found to provide reasonable predictive fit of individual demand. Similar to the logit model, this approach retains the link to consumer theory, and at the same time avoids the specification bias of the logit. In our implementation, we assume that variation in quantity choices reflects variation in consumption. We do not model the potential dynamics that arise when consumers stock-pile in anticipation of future price increases. While advances have been made in addressing such dynamics in the context of household data (Erdem et al. 2003, Hendel and Nevo 2003), incorporating such sophisticated choice behavior at the aggregate level is beyond the scope of the current analysis.

We show how the structural parameters of the derived demand system can be estimated with readily available aggregate store-level data. Our specification allows us to control for the role of unobserved heterogeneity in consumer tastes as well as the potential endogeneity of prices, both of which could bias our parameter estimates. The endogeneity of prices results from the presence of unobserved brand characteristics that could influence consumer choices and that could correlate with prices (Besanko et al. 1998). We control for this problem using an instrumental variables procedure similar in spirit to the approach used for the logit demand system (Berry et al. 1995). Specifically, we propose a modified version of the contraction-mapping found in Berry et al. (1995) to “invert” the mean utility out of the model. The mean utilities are then used to construct moment conditions that form the basis of a method of simulated moments estimation procedure.

We conduct several simulation experiments to show the proposed model’s parameters are identified and that the specification is more robust to quantity choice behavior than the typical logit demand system. First, we simulate data from the proposed discrete/continuous demand system and we find we are

able to recover the structural parameters fairly accurately. In contrast, fitting a logit model to such data not only fails to recover the structural parameters, it also provides incorrect estimates of the price elasticities. Hence, the logit does not appear to be a good “approximation” of demand when quantity choices matter. We next simulate data from the logit demand system. In this case, we find that the logit model, as expected, recovers structural parameters and provides reasonable estimates of price elasticities. At the same time, our proposed discrete/continuous demand system also recovers comparable estimates of the price elasticities. We conclude that our proposed approach serves as a more robust demand specification for packaged-good categories as it can accommodate substitution patterns across a wider scope of quantity choice behaviors.

The proposed model is estimated using weekly refrigerated orange juice data for 30 stores from a large Chicago supermarket chain. We compare the results from our proposed model (purchase incidence, brand choice, and quantity model that accounts for primary and secondary demand effects) with those from a logit (purchase incidence and brand-choice-only model). We begin by using a category pricing model to compute margins for each demand specification. Since wholesale prices are observed in our data, we know the true retail margins for the brands in the product category. We test which of the demand models better predicts the true observed margins in the data. We find that the proposed model performs better than the logit model in predicting the observed margins both in and out of sample.

We then contrast the two models’ substantive implications in terms of price elasticities. We find that the elasticity estimates from the proposed model are significantly lower than those of the logit. The logit predicts conditional purchase elasticities (i.e., brand-switching effects) about 56% higher and category purchase probabilities (i.e., category expansion effects) about 82% higher than our proposed model on average. Both these effects are consequences of the single-unit assumption. We then decompose price effects into primary (purchase incidence/quantity) and secondary (brand-switching) demand effects using an elasticity-based decomposition (Bell et al. 1999) and a unit sales-based decomposition (van Heerde et al. 2003). This comparison provides some external validity to our approach since our results, based on aggregate data, are consistent with those reported in previous research, using household data. While the use of aggregate data is similar to van Heerde et al. (2004), our model formulation allows us to derive a formal decomposition of the price elasticity of demand. Next, we use the two models to predict the impact of a hypothetical 30 cent price cut

³ A separate literature has explored the more sophisticated case when consumers purchase assortments consisting of multiple brands and multiple unit quantities (Dubé 2004, Kim et al. 2002).

on the store brand, potentially to generate new trials. Since the logit model a priori constrains consumers to single-unit purchases, it predicts a much higher category expansion effect than the proposed model. Taken together, these results favor our proposed model in capturing the true nature of demand in this category.

2. Model Formulation

Our main objective is to derive an aggregate demand system that provides flexible substitution patterns between brands while still allowing for both primary and secondary demand considerations at the consumer level. We begin this section with the derivation of such a consumer-level model. Then, we derive the corresponding aggregate demand system.

2.1. Consumer-Level Model

In this section, we outline the consumer-level models used by Chiang (1991) and Chintagunta (1993), which are based on that of Hanemann (1984). In the following section, we derive the corresponding aggregate demand system. A utility-maximizing consumer chooses one (and only one) of the J brands from a product category in store s during week t . The consumer may purchase a variable quantity of the chosen brand. The consumer may also choose not to buy any of the products, in which case the entire shopping budget is allocated to a numeraire capturing expenditures on other items in the store outside the category of interest.

Formally, consumer i chooses the utility-maximizing bundle, subject to a budget constraint:

$$\begin{aligned} \max_{x_{i1st}, \dots, x_{ijst}} \quad & u = u^* \left(\sum_{j=1}^J \psi_{ijst} x_{ijst}, \psi_{izst} z_{ist} \right) \\ \text{s.t.} \quad & \sum_{j=1}^J p_{jst} x_{ijst} + z_{ist} = y_{ist}; \quad x_{ijst} \geq 0; \quad z_{ist} > 0, \end{aligned} \quad (1)$$

where x_{ijst} is the quantity purchased of brand j ($j = 1, \dots, J$), ψ_{ijst} is the perceived quality-index of brand j by consumer i , z_{ist} is the numeraire good, ψ_{izst} is consumer i 's perceived quality index of the outside good, p_{jst} is the price for brand j faced by the consumer in store s in week t , y_{ist} is the total expenditure outlay by the consumer. The linear subutility over the J brands ensures that only a single alternative is chosen. Analogous to Chiang (1991) and Chintagunta (1993), we assume the following functional form for the perceived quality index of each brand and the outside good:

$$\begin{aligned} \psi_{ijst} &= \exp \left[\frac{1}{\alpha_s} (\gamma_{ijst} + \beta_{ist} d_{jst} + \xi_{jst} + \varepsilon_{ijst}) \right], \\ \psi_{izst} &= \exp(\varepsilon_{izst} / \alpha_s). \end{aligned} \quad (2)$$

Here α_s is a store-specific scale that shifts the perceived quality of brands and the outside good across stores; γ_{ijst} is the intrinsic taste for brand j for a consumer i in store-week st ; and d_{jst} is a deal variable that indicates an in-store display for brand j in store-week st with corresponding consumer sensitivity parameter β_{ist} . The random disturbances, ε_{ijst} and ε_{izst} , are consumer- and alternative-specific unobserved factors that affect the consumer's valuation of brand j and the outside alternative in store-week st respectively. We assume the unobserved components of the perceived qualities are independent and identically distributed extreme value such that $\varepsilon_{ist} = (\varepsilon_{i1st}, \dots, \varepsilon_{ijst}) \sim \text{EV}(0, \mu)$ and $\varepsilon_{izst} \sim \text{EV}(D_s' \tau, \mu)$, where D_s is a vector of demographic variables specific to store s , and τ is a vector of parameters to be estimated.⁴ By allowing store demographics to affect the distribution of the perceived quality of the outside good, we allow the relative attraction of purchasing in the category to vary across stores. In contrast to previous work with this model, we also include the term ξ_{jst} . The error component ξ controls for additional unobserved (to the researcher) characteristics specific to brand j and store-week st , such as shelf-space allocation and positioning, that influence consumer choices. In the estimation section we discuss the econometric challenges that arise if this term is correlated with marketing variables, such as shelf prices.

Finally, as in Chiang (1991), we assume the indirect utility function corresponding to (1) has the flexible Homothetic Indirect TransLog (HITL) form. The solution to (1) implies that the indirect utility is a function of the expenditure outlay, y_{st} , and the "quality-weighted" prices of the chosen brand, p_{jst}/ψ_{ijst} , and the numeraire good, $1/\psi_{izst}$. Using Roy's identity, the demand function (conditional on category purchase, $I_{ist} = 1$, and choice of brand j , $C_{ijst} = 1$) corresponding to the HITL is:

$$\begin{aligned} x_{ijst}(p_{jst}, \psi_{ijst}, \psi_{izst}, y_{st} | C_{ijst} = 1, I_{ist} = 1) \\ = (y_{st}/p_{jst}) [\alpha_1 - \alpha_3 \ln(p_{jst}/\psi_{ijst}) + \alpha_3 \ln(1/\psi_{izst})], \end{aligned} \quad (3)$$

where α_1 and α_3 are parameters of the HITL indirect utility function (for further details on the HITL, see Pollak and Wales 1992). These assumptions give rise to the expected conditional demands

$$\begin{aligned} E_e(x_{ijst}) &= \int_{j \text{ chosen}} x_{ijst}(p_{jst}, \psi_{ijst}, \psi_{izst}, y_{st} | C_{ijst} = 1, I_{ist} = 1) d\varepsilon_{ist} \\ &= -\frac{1}{\alpha_s / \mu \alpha_3 p_{jst}} \left[\frac{\ln\{P(I_{ist} = 0)\}}{P(I_{ist} = 1)} \right] \end{aligned} \quad (4)$$

⁴ Note that for a given store, the demographic variables will be the same for all weeks. Additionally, in our data, we have access to only the mean demographics for each store—this implies that for a given store, the demographic variables will be the same for all individuals in each week. Hence we drop the subscripts i for individual and t for week for the demographic variables.

and corresponding choice probabilities

$$\Pr(I_{ist} = 1) = \sum_{j=1}^J e^{V_{ijst}} / \left(1 + \sum_{j=1}^J e^{V_{ijst}} \right); \quad (5)$$

$$\Pr(C_{ijst} = 1, I_{ist} = 1) = e^{V_{ijst}} / \left(1 + \sum_{j=1}^J e^{V_{ijst}} \right),$$

where $V_{ijst} = [\gamma_{ijst} + \alpha_1 \alpha_s / \alpha_3 + \beta_{ist} d_{jst} - \alpha_s \ln(p_{jst}) - D'_s \tau + \xi_{jst}] / \mu$ (derivations of all equations are available from the authors). Equations 4 and 5 describe the quantity and incidence/brand choice components of the model, both of which are derived from the same underlying utility maximization problem (1).

To control for heterogeneity across consumers shopping during store-week st , we include random coefficients in the perceived quality function, ψ_{ijst} (Arora et al. 1998). To control for differences in mean effects across stores, we interact taste parameters with store-specific demographic variables, D_s^* (note that D_s^* could contain the same or different demographic variables as D_s). We account for heterogeneity in the following way:⁵

$$(\gamma_{i1st}, \dots, \gamma_{ijst}, \dots, \gamma_{iJst})' \\ = (\gamma_1, \dots, \gamma_j, \dots, \gamma_J)' + L \boldsymbol{\varpi}_{ist} \quad (6.1)$$

such that $\boldsymbol{\varpi}_{ist} \sim MVN(0, I_{J \times 1})$ and $L'L = \Sigma_{J \times J}$, and,

$$\beta_{ist} = \beta + \sigma_d s_{ist}, \quad s_{ist} \sim N(0, 1) \\ \alpha_s = \bar{\alpha} + D_s^* \theta \quad (6.2)$$

This formulation of the demand model differs from previous work in two ways. First, we include the error component ξ to control for unobserved attributes. We also include interactions of preferences with demographics to control for differences in demand across stores.

2.2. Aggregate Demand

We now derive the aggregate demand system corresponding to the consumer model of the previous section. We define the potential market size as the mass of consumers who shop in the store s during week t , N_{st} . For the current analysis, we assume that N_{st} is exogenous. Hence, variation in prices of the various orange juice SKUs are assumed to have no impact on total store traffic, only on category size (share of store traffic that purchases in the category). This assumption could be problematic in store traffic-generating categories such as carbonated soft drinks.⁶

⁵ Though not shown, it is straightforward to include store-demographics effects for the brand-intercepts and for the effects of the brand-attributes (e.g., Chintagunta et al. 2003).

⁶ Although not reported, we found that observed variation in store traffic does not co-vary significantly with an orange juice price index, whereas the former does correlate with indices from other categories.

To derive aggregate demand for brand j , we integrate over the set of consumers choosing brand j :

$$Q_{jst} = N_{st} \int [\Pr(C_{ijst} = 1, I_{ist} = 1) E_\epsilon(x_{ist})] \phi(\Lambda) \partial \Lambda, \quad (7)$$

where $\Lambda = (\boldsymbol{\varpi}, s)'$ and $\phi(\cdot)$ denote the pdf of a standard multivariate normal distribution. The corresponding average quantity purchased per customer, $\tilde{Q}_{jst} \equiv Q_{jst} / N_{st}$, is obtained from (7):

$$\tilde{Q}_{jst} = \int [\Pr(C_{ijst} = 1, I_{ist} = 1) E_\epsilon(x_{ist})] \phi(\Lambda) \partial \Lambda \quad (8)$$

To assess the importance of modeling purchase quantity considerations, we compare the above aggregate demand system with the logit aggregate demand system, where the latter does not account for quantity choices. In the case of the logit demand system, the average quantity purchased per customer (i.e., the market share) can be written as:

$$\tilde{Q}_{jst} = \int \Pr(C_{ijst} = 1, I_{it} = 1) \phi(\Lambda) \partial \Lambda. \quad (9)$$

Essentially, (9) is equivalent to imposing $E_\epsilon(x_{ist}) = 1$ in (8). Empirically, we can compare parameter estimates and marginal effects under (8) and (9) to measure the role of purchase quantity considerations.

3. Model Estimation

We now outline the method-of-simulated-moments (MSM) procedure that produces consistent estimates of the model parameters. The procedure also resolves a potential endogeneity problem inherent in the model due to the possible correlation between observed prices and characteristics specific to the brands in each store-week that are unobserved to the researcher, ξ_{jst} .

Combining equations (4), (5), (6), and (8), we rewrite the aggregate demand for brand j as:

$$\tilde{Q}_{jst}(\Gamma) = \int \frac{y_{st}}{\alpha_s p_{jst}} \frac{e^{V_{ijst}}}{\sum_{k=1}^J e^{V_{ikst}}} \\ \times \ln \left[1 + \sum_{k=1}^J e^{V_{ikst}} \right] \phi(\Lambda) \partial \Lambda, \quad (10)$$

where

$$V_{ijst} = [(\gamma_j + \bar{\alpha} \alpha_1 / \alpha_3) - \alpha_s \ln(p_{jst}) + d_{jst} \beta \\ + D_s^* \theta \alpha_1 / \alpha_3 - D'_s \tau + \xi_{jst}] \frac{1}{\mu} \\ + [L \boldsymbol{\varpi}_{ist} + d_{jst} \sigma_d s_{ist}] \frac{1}{\mu}$$

and Γ is a vector containing all the model parameters. In principle, one could estimate (10) using maximum likelihood, where ξ is the econometric error term. The

concern is that the error term captures unobserved (to the researcher) demand-shifting factors that may be observed by the retailer. If the retailer accounts for these factors when setting prices, then ξ would be correlated with prices, which would in turn generate an endogeneity bias. This type of bias has been discussed and documented in the context of similar weekly retail data (Besanko et al. 1998, Sudhir 2001, Chintagunta et al. 2003). To resolve the endogeneity problem, we rewrite V_{ijst} in (10) as follows:

$$\tilde{Q}_{jst}(\delta; \Theta) = \int \frac{y_{st}}{\alpha_s p_{jst}} \frac{e^{\delta_{jst} + \Omega_{ijst}}}{\sum_{k=1}^J e^{\delta_{kst} + \Omega_{ikst}}} \cdot \ln \left[1 + \sum_{k=1}^J e^{\delta_{kst} + \Omega_{ikst}} \right] dF(\Omega), \quad (11)$$

where $\delta_{jst} = [(\gamma_j + \bar{\alpha}\alpha_1/\alpha_3) - \alpha_s \ln(p_{jst}) + d_{jst}\beta + D_s^* \theta \alpha_1/\alpha_3 - D_s^* \tau + \xi_{jst}]/(1/\mu)$ is the component of V_{ijst} that is common across consumers in store-week st , and $\Omega_{ijst} = (L\varpi_{ist} + d_{jst}\sigma_d s_{ist})/(1/\mu)$ is the consumer-specific component. Due to linearity, we can write δ_{jst} as $\delta_{jst} = X_{jst}B - \alpha_s \ln(p_{jst}) + \xi_{jst}$, where B contains an intercept term, $\gamma_j/\mu + \bar{\alpha}\alpha_1/\alpha_3\mu$; a mean sensitivity to display, β/μ ; and sensitivities to demographic variables. The scale μ cannot be recovered separately and is normalized to 1. In addition, we normalize α_3 to 1 since it cannot be identified separately from α_1 . The term $\alpha_s (= \bar{\alpha} + D_s^* \theta)$ is henceforth referred to as the price sensitivity for store s . The consumer-specific term, Ω_{ijst} , depends on the remaining parameters, L, σ_d . Denote the entire set of parameters to be estimated as $\Theta = \{B, \bar{\alpha}, \theta, L, \sigma_d\}$.

Analogous to Berry (1994), we invert (11) to recover the vector $\delta_{jst}(\Theta)$. In the appendix we prove that (11) is indeed an invertible function of δ_{jst} . Hence, for a given set of model parameters and given values for ξ_{jst} , there exists a *unique* vector δ_{jst} such that (11) holds identically (i.e., $q_{jst} = \tilde{Q}_{jst}(\delta; \Theta)$). To invert δ_{jst} out of the model, we use a modified version of the contraction-mapping proposed by Berry et al. (1995) for the logit demand system (in the appendix, we prove that the contraction-mapping is valid in the context of our demand specification). We denote the average quantity per customer in the aggregate data as q_{jst} . Now define the function $g(\cdot): \mathbb{R}^J \rightarrow \mathbb{R}^J$ as:

$$g(\delta) = \delta + \ln(q) - \ln[\tilde{Q}(\delta)]. \quad (12)$$

For each guess of the parameters Θ , we iterate on (12) to recover the unique vector $\delta_{jst}(\Theta)$ that solves (11) across all store-weeks. Intuitively, we calibrate values of $\delta_{jst}(\Theta)$ that exactly fit the predicted \tilde{Q}_{jst} to the observed q_{jst} in the data. To evaluate the multidimensional integrals in (11), we use Monte Carlo simulation. Thus, the inversion procedure described above

matches the *simulated* \tilde{Q}_{jst} to the observed average q_{jst} in the data. Simulation was carried out using 100 draws.⁷

Using our values of $\delta_{jst}(\Theta)$, we construct moment conditions based on $\xi_{jst} = \delta_{jst}(\Theta) - X_{jst}B + \alpha_s \ln(p_{jst})$, where we assume $E[\xi_{jst} X_{jst} | X_{jst}] = 0$. The endogeneity of prices arises if $E[\xi \ln(p) | \ln(p)] \neq 0$. Hence, in addition to X_{jst} , we need additional instruments, Z_{jst} , such that $E[\xi_{jst} Z_{jst} | Z_{jst}] = 0$. Using these orthogonality conditions, we can estimate our model parameters consistently using a standard method of simulated moments procedure (Pakes and Pollard 1989). Further technical details of the procedure are provided in the appendix.

3.1. Simulation Studies

We conduct several simulation experiments to illustrate the identification of our proposed model's parameters and to demonstrate its robustness across alternative quantity choice behavior at the individual level. For now we summarize our findings, referring the reader to the appendix for complete results and details. First, we simulate data from the proposed discrete/continuous demand system and find we are able to recover the structural parameters fairly accurately. In contrast, fitting a logit model to such data not only fails to recover the structural parameters, it also provides incorrect estimates of the price elasticities. In particular, conditional on purchase, when quantities at the individual-level range from 2–4 units, the logit overstates aggregate brand-choice elasticities by about 44%, and aggregate category purchase elasticities by about 31% on average. By construction, all variation in market-shares has to reflect changes in brand-switching or purchase incidence in the logit; and consequently, it overstates the corresponding elasticities when the share-variation is also driven by changes in underlying quantity choices. Hence, the logit does not appear to be a good “approximation” of demand when quantity choices matter. We next simulate data from the logit demand system. In this case, we find that the logit model, as expected, recovers structural parameters and provides reasonable estimates of price elasticities. At the same time, our proposed discrete/continuous demand system also recovers comparable estimates of the price elasticities. Finally, we simulate data from a household level model where quantity choices are discrete. A potential limitation of our proposed model is that quantities are treated as divisible and, hence, we wish to see if the continuous approximation of quantity choices can still provide meaningful demand estimates. We use a brand-choice-then-quantity model in which the

⁷ We found little impact increasing the number beyond 100 to 150 or 200 draws.

probability of purchase is given by a logit and the discrete quantity chosen is given by a truncated Poisson. Again, we find that our proposed model provides fairly reasonable estimates of the underlying choice probabilities and conditional quantities. Hence, we conclude that our proposed approach provides a more robust demand specification for packaged-good categories as it can accommodate substitution patterns across a wider scope of quantity choice behaviors.

4. Data and Results

4.1. Data

We use aggregate store-level data for the refrigerated orange juice category obtained from the Dominick's Finer Foods (DFF) database at the University of Chicago. The IRI factbook documents that in this category consumers typically purchase 5.47 pints (109.4 oz.) of a single brand of refrigerated juice on a given purchase occasion. Using the most common pack size in this category, 64 oz., this corresponds to an average demand of 1.71 units per purchase occasion (in the Dominick's dataset, products with pack-size of 64 oz. had around 88% of the market share; see Table 3). Using a separate household database, from a different market (Denver), we find that, conditional on the purchase of refrigerated orange juice, 99% of the trips involve the purchase of a single SKU, but over 20% of the trips involve the purchase of multiple units of an SKU.

Our data consist of weekly sales, prices, displays, and profit margins at the UPC-level along with total weekly store traffic for 30 randomly-selected stores during the 52-week period of 1992. We also use the corresponding mean demographic variables for each store. Descriptions of these demographic variables can be found in Chintagunta et al. (2003). We focus our analysis on the top 7 SKUs (we combined UPCs of the same brand and package size whenever the correlation in their prices exceeds 0.8 across stores and weeks). In Table 1, we present descriptive statistics of these data. Since our data contain two different pack-sizes, 64 oz. and 96 oz., we include alternative specific intercepts in the model to control for volume differences. Hence, in predicting quantity choices, the model controls for the difference in volume between, for instance, 2 units of a 64-oz. product versus 2 units of a 96-oz. product (Allenby et al. 2004). As in Chintagunta et al. (2003), we use weekly wholesale prices as additional instruments for shelf prices. The motivation for using wholesale prices as instruments comes from the interpretation of ξ_{sjt} as store-specific aggregate shocks (i.e., aggregate shocks for all consumers shopping in a given store). We refer

Table 1 Product-Level Averages Across Store-Weeks

Brand	Size (oz.)	Share (%)	Price	Cost	Display
Minute Maid	64	21.25	\$2.25	\$1.69	0.11
Minute Maid	96	4.35	\$4.11	\$2.98	0.03
Dominick's	64	22.71	\$1.65	\$1.15	0.01
Tropicana Prm	64	21.87	\$2.67	\$1.88	0.09
Tropicana SB	64	17.69	\$2.38	\$1.62	0.05
Tropicana Prm	96	8.12	\$4.52	\$3.26	0.07
Florida	96	3.97	\$2.18	\$1.48	0.04

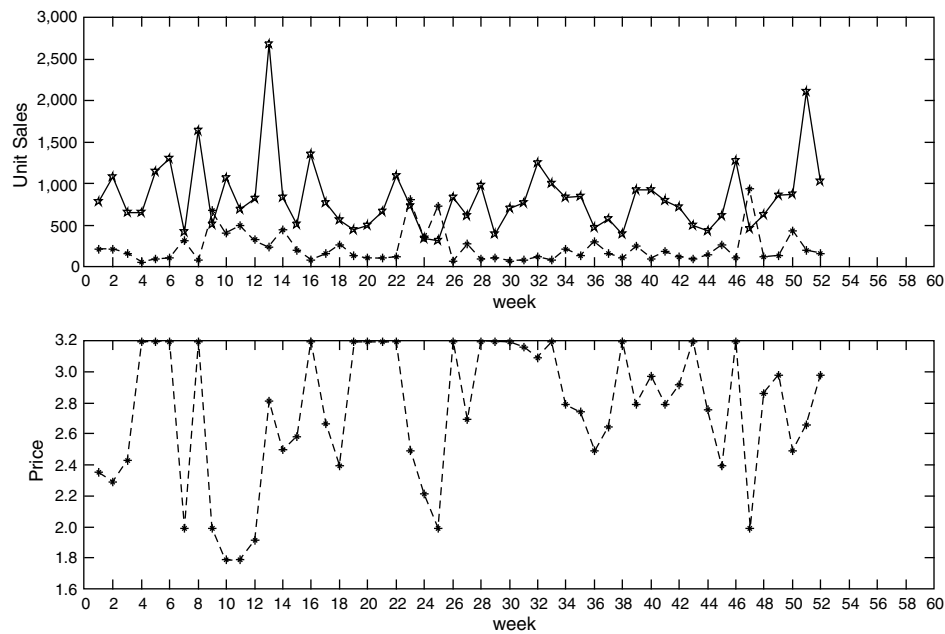
the reader to Chintagunta et al. (2003) for a discussion. A limitation of this approach is that it would not resolve endogeneity arising from more macroeconomic shocks, such as unobserved (to the researcher) television advertising, which would most likely be common across all stores and, moreover, would be correlated with wholesale prices. The R^2 for the first-stage regression of prices on the instruments is 0.728.⁸ We use weekly store-traffic as the potential market size for store s during week t . Finally, we operationalize y_{st} by dividing the total dollar sales for the grocery section in a store-week by the corresponding level of store-traffic.

In our raw data, we do observe patterns consistent with brand-switching as well as purchase acceleration. In Figure 1, we plot the unit sales and prices of Tropicana Premium (dotted line with asterisk) against all other brands (solid line) in our sample Store 2. The bottom panel shows the per-unit price of Tropicana Premium in Store 2. During weeks 8 and 9, we observe quantity changes consistent with brand switching. The price of Tropicana Premium dropped from \$3.2 to \$2.0 and its sales increased by around 700 units, while those of all the other brands dropped by around 1000 units. During other promotion weeks, however, we see similar price-cut induced sales increases for Tropicana without any effect on sales of other brands (e.g., weeks 24 and 25), possibly due to either category expansion or purchase acceleration (Bell et al. 1999, van Heerde et al. 2003).

4.2. Results

We now report the results of estimation of the proposed model. Preliminary analysis indicates that ignoring either heterogeneity or endogeneity results in a significantly lower magnitude for the estimated price parameter, a finding consistent with previous literature. We only present the results from the model

⁸ The full set of instruments includes all demand-shifting variables other than price along with wholesale prices. A potential concern is that some of the marketing variables, displays, are also correlated with ξ . Previous research has routinely rejected the hypothesis that displays correlate with ξ and, hence, are invalid instruments (e.g., Sudhir 2001, Chintagunta et al. 2003). A Hausman test supports the exogeneity assumption of displays for our data.

Figure 1 Unit Sales for Tropicana Premium and All Other Brands (Top) and Prices for Tropicana Premium (Bottom), in Store 2

that allows for heterogeneity and accounts for price endogeneity (Table 2). To allow demand to vary across stores, we include demographic variables as follows:

$$\alpha_s = \bar{\alpha} + \theta_{\text{ethnic}} \text{ethnic}_s + \theta_{\text{hval1500}} \text{hval150}_s,$$

$$D_s \tau = \tau_{\text{drvtm}} \text{drvtm}_s + \tau_{\text{age60}} \text{age60}_s + \tau_{\text{hhlarge}} \text{hhlarge}_s,$$

where α_s is the price parameter. An unrestricted variance-covariance matrix for the intercepts, Σ , is estimated.

Looking at the results in Table 2, we find that consumers have a higher preference for Tropicana Season's Best and Tropicana Premium versus Minute Maid and the store brand. As expected, displays have a significant positive effect on the utilities of all the inside goods, and prices have a significant negative effect. We do not find much unobserved heterogeneity in the effects of display across stores ($\sigma_d = 0.103$). Heterogeneity in the price effect is captured by interacting the (log of the) price variable with store demographics. We find that a higher proportion of African-American or Hispanic families in the store area increases the mean price sensitivity of consumers in that store. This is as expected, since African-American and Hispanic-dominated areas in our data tend to have lower incomes on average. We find that consumers are less price sensitive in store areas with a higher proportion of households with properties valued greater than \$150,000. Although not reported, we find evidence of substantial differences in the mean price sensitivity across stores, driven mainly by the significant demographic effects and interactions. This finding is consistent with previous

research using the DFF data (Hoch et al. 1995 and Chintagunta et al. 2003).

In Table 3, we present the predicted average quantity per customer conditional on purchase, $E_e \tilde{x}_{ijst}(p_{jst}, \psi_{ijst}, \psi_{izst}, y_{st} | C_{ijst} = 1, I_{ist} = 1)$, across all store-weeks. The results indicate that, on average, consumers tend to purchase more units of the store brand than the other brands. This result arises since, conditional

Table 2 Parameter Estimates (with Heterogeneity)

	Parameter	<i>t</i> -Stat	Standard deviations of intercepts	
			Parameter	<i>t</i> -Stat
Minute Maid 64 oz.	−19.278	−2.733	8.145	3.159
Minute Maid 96 oz.	−5.676	−5.723	2.905	5.831
Dominick's 64 oz.	−14.250	−3.173	7.474	3.784
Tropicana Prm 64 oz.	−7.491	−5.331	5.251	7.152
Tropicana SB 64 oz.	−5.207	−6.316	2.320	3.483
Tropicana Prm 96 oz.	−4.530	−4.413	1.806	2.844
Florida 96 oz.	−5.684	−9.016		
-Log(price)	2.336	13.839		
Display	0.387	3.626		
-Log(price)*ethnic	4.165	5.083		
-Log(price)*hval150	−1.363	−6.175		
Ethnic	8.750	6.505		
Hval150	−1.669	−3.981		
Drvtm	−0.051	−2.514		
Age60	1.367	3.396		
Hhlarge	0.013	0.015		
σ_d	0.103	0.094		
GMM objective value			395.614	

Notes. Parameters represent V_{jst} in Equation 14; total parameters = 45; observations = 10,682 (product-store-week combinations); full variance-covariance matrix for intercepts is estimated; *t*-statistics of intercept heterogeneity terms were calculated using 500 bootstrap draws of parameter vector.

Table 3 Predicted Expected Quantities Purchased Across Store-Weeks

Product	Mean	Std. deviation
Minute Maid 64 oz.	1.715	0.625
Minute Maid 96 oz.	0.932	0.323
Dominick's 64 oz.	2.352	0.920
Tropicana Prm 64 oz.	1.450	0.555
Tropicana SB 64 oz.	1.651	0.670
Tropicana Prm 96 oz.	0.839	0.280
Florida 96 oz.	1.751	0.633

on brand-choice, the predicted quantity purchased is proportional to the ratio expenditure/price. Since the store brand has the lowest price per unit (\$1.65 on average), the model predicts that consumers buy it in the highest quantities.

5. Model Implications

We now discuss the substantive implications of the model and compare it along several dimensions to the popular random coefficients logit demand system, which assumes single-unit purchases at the consumer level. First, we compare the two specifications in their ability to predict the observed markups (price – wholesale costs) in the data, both in and out-of-sample. Here we show how we can test model superiority in terms of predictive fit. Subsequently, we compare the models in terms of their predictions for price elasticities of demand and purchase incidence. We also compare their predictions for a hypothetical price cut. Although not reported, we also found considerable improvement in fit of our proposed model relative to a log-log demand specification.

5.1. Margin Predictions

We propose a simple statistical test for the relative superiority of the proposed model versus the logit. The wholesale prices in the database enable us to compute the true retail margins across the store-weeks in our sample. For each of the demand models, we then use a category pricing model to obtain predicted margins. The category pricing model has been found to provide a reasonable approximation of pricing in supermarket categories (Chintagunta et al. 2003). In Table 4, we find the model that accounts for primary and secondary demand effects provides predicted margins that are more highly-correlated with the true margins and are, on average, more accurate.

Following Chintagunta et al. (2003), we also test which model provides a better fit of the observed margins using the following minimum distance metric:

$$\min_{\Delta} (\text{margin} - \Delta)' \hat{\Phi} (\text{margin} - \Delta).$$

The term Δ is the estimated margin and $\hat{\Phi}$ is the covariance matrix of the observed margins. The minimum distance criterion corresponding to both models are presented in the last row of Table 4. The fit for the single-unit logit model is very similar to that found by Chintagunta et al. (2003) using DFF data for the same category. However, the proposed model provides a better fit according to the minimum distance criterion. For robustness, we repeat the minimum distance procedure using margins from the 53 remaining stores, not used at the demand estimation stage. We compute the margins for each brand in each of 2,724 store-weeks from our hold-out sample.⁹ The results are presented in Table 5. We see that the proposed model once again performs better than the single-unit model out-of-sample.

5.2. Price Elasticities

In Tables 6 and 7 we present the estimated own and cross-price elasticities for the proposed and the logit models respectively. For the former, we decompose the elasticity into an unconditional brand choice elasticity and an expected conditional quantity elasticity. The unconditional brand choice elasticity is computed as the mean of

$$\varepsilon p_{p_k}^j = p_{kst} \left(\int \partial P(C_{ijst} = 1, I_{ist} = 1) / \partial p_{kst} \phi(\Lambda) \partial \Lambda \right) / \left(\int P(C_{ijst} = 1, I_{ist} = 1) \phi(\Lambda) \partial \Lambda \right),$$

and the expected conditional quantity elasticity is computed as the mean of:

$$\varepsilon q_{p_j}^j = p_{jst} \left(\int \partial E_e(x_{ijst} | C_{ijst} = 1, I_{ist} = 1) / \partial p_{jst} \phi(\Lambda) \partial \Lambda \right) / \left(\int E_e(x_{ijst} | C_{ijst} = 1, I_{ist} = 1) \phi(\Lambda) \partial \Lambda \right)$$

across all store-weeks in the sample. As in the case of a logit model, the inclusion of heterogeneity provides more flexible substitution patterns.¹⁰

The net elasticity estimates, the sum of brand-choice and quantity elasticities, from the proposed model are presented in Table 6a. We find that Minute Maid 64 oz. has the least price-elastic demand, while Florida Orange has the most price-elastic demand. Consistent with previous studies (Blattberg and Wisniewski 1989, Allenby and Rossi 1991), we find

⁹ In predicting margins out-of-sample, it is important to control for the structural error ξ_{jst} . From our estimation, for each week, we have estimates of ξ_{jst} for the 30 stores in our sample. When computing margins for a store-week out-of-sample, we integrate over the empirical distribution of ξ across the 30 DFF stores for that week.

¹⁰ With no heterogeneity, the own-price elasticities are roughly proportional to average unconditional quantities, and cross-price elasticities are independent of own brand prices/attributes.

Table 4 Sample Prediction of Margins

Brand	Size (oz.)	Conditional share (%)	Average price	Average cost	True % margins	Logit model		Proposed model	
						Predicted % margins ¹	Correlations of true and pred: margin ²	Predicted % margins	Correlations of true and pred: margins
Minute Maid	64	21.25	\$2.25	\$1.69	23.83	50.90	0.3688	41.99	0.5757
Minute Maid	96	4.35	\$4.11	\$2.98	26.42	38.43	0.4583	33.60	0.6214
Dominick's	64	22.71	\$1.65	\$1.15	28.41	49.93	0.3794	40.56	0.6360
Tropicana Prm	64	21.87	\$2.67	\$1.88	28.41	53.84	0.3432	36.00	0.5737
Tropicana SB	64	17.69	\$2.38	\$1.62	30.39	52.13	0.4223	34.46	0.6954
Tropicana Prm	96	8.12	\$4.52	\$3.26	27.20	51.70	0.2320	36.47	0.5986
Florida	96	3.97	\$2.18	\$1.48	31.80	31.94	0.4775	32.44	0.5021
Minimum distance criterion						1.3068		0.2991	

¹ "% margin" = $100 * (p - c) / p$.² "margin" = $(p - c)$.**Table 5** Out of Sample Prediction of Margins

Brand	Size (oz.)	Conditional share (%)	Average price	Average cost	True % margins	Logit model		Proposed model	
						Predicted % margins	Correlations of true and pred: margin	Predicted % margins	Correlations of true and pred: margins
Minute Maid	64	21.16	\$2.23	\$1.69	23.61	48.87	0.2403	49.53	0.3894
Minute Maid	96	4.02	\$4.08	\$2.99	26.03	36.98	0.2022	31.48	0.3621
Dominick's	64	25.33	\$1.64	\$1.15	28.03	48.26	0.2885	45.33	0.5237
Tropicana Prm	64	19.75	\$2.66	\$1.88	27.95	51.73	0.2439	35.81	0.4385
Tropicana SB	64	17.75	\$2.36	\$1.62	29.86	49.99	0.3088	31.67	0.5636
Tropicana Prm	96	7.47	\$4.50	\$3.26	26.94	55.18	0.0658	34.42	0.3355
Florida	96	4.53	\$2.17	\$1.48	31.41	29.30	0.2693	29.16	0.3053
Minimum distance criterion						1.2975		0.3083	

Table 6(a) Expected Price-Elasticities of Demand (Across Stores-Weeks) for the Proposed Model

Effect of change in price of:	On the demand for:						
	MM64	MM96	Dom64	TRPrm64	TRSB64	TRPrm96	FLR
MM64	−2.6274*	0.0059	0.3610*	0.0058	0.0535	0.1310*	0.0007
MM96	0.0033	−3.0632*	0.0088	0.0122	0.0073	0.0063	0.0053
Dom64	0.3480*	0.0145	−2.7468*	0.0022	0.0340	0.0637*	0.0013
TRPrm64	0.0091	0.0354*	0.0034	−2.7503*	0.0341	0.0396*	0.0078
TRSB64	0.1010*	0.0208	0.0698	0.0355	−3.1396*	0.1190*	0.0051
TRPrm96	0.1450*	0.0113*	0.0783*	0.0274*	0.0783*	−3.0359*	0.0015
FLR	0.0003	0.0034*	0.0005	0.0017	0.0011	0.0006	−3.1920*

* Statistically significantly different from zero at the 10% level.

Table 6(b) Expected Unconditional Choice Price Elasticities (Across Stores-Weeks) for the Proposed Model

Effect of change in price of:	On the unconditional probability of choice of:							No purchase probability
	MM64	MM96	Dom64	TRPrm64	TRSB64	TRPrm96	FLR	
MM64	−1.6200	0.0059	0.3610	0.0058	0.0535	0.1310	0.0007	0.0050
MM96	0.0033	−2.0600	0.0088	0.0122	0.0073	0.0063	0.0053	0.0057
Dom64	0.3480	0.0145	−1.7400	0.0022	0.0340	0.0637	0.0013	0.0056
TRPrm64	0.0091	0.0354	0.0034	−1.7400	0.0341	0.0396	0.0078	0.0133
TRSB64	0.1010	0.0208	0.0698	0.0355	−2.1300	0.1190	0.0051	0.0157
TRPrm96	0.1450	0.0113	0.0783	0.0274	0.0783	−2.0300	0.0015	0.0086
FLR	0.0003	0.0034	0.0005	0.0017	0.0011	0.0006	−2.1900	0.0040

Table 6(c) Expected Conditional Quantity Elasticities (Across Stores and Weeks) for the Proposed Model

MM64	−1.0074
MM96	−1.0032
Dom64	−1.0068
TRPrm64	−1.0103
TRSB64	−1.0096
TRPrm96	−1.0059
FLR	−1.0020

evidence for asymmetric switching between national and store brands, with the national brands gaining a bit more share from the store brand from its price-cuts, than vice versa. We also observe high cross-price elasticities for different pack-sizes of the same brand. The patterns in the unconditional choice elasticities (Table 6b) reflect those in the demand elasticities (Table 6a). Similar to previous research with household data, we find quantity elasticities in Table 6c close to 1 (Chiang 1991 and Chintagunta 1993).¹¹

The demand elasticities for the logit model are presented in Table 7. Under the single-unit purchase assumption, any change in a brand's market share must reflect either brand-switching or purchase incidence. Consequently, we expect the model would *overstate* the unconditional brand-choice probability if consumers do indeed respond to price changes by changing their quantity choices. Comparing Tables 6b and 7, we observe that the unconditional choice probability elasticities for the logit model are systematically higher than those of the proposed model. Although not reported, mean conditional brand choice elasticities differ by roughly 56%. Despite similarities in the mean own-price elasticities of demand for the two models, we find considerable differences when we look at individual store-weeks. We find that own-price elasticities from the two models can differ by as much as 50% in a given store-week.

To test whether the total demand elasticities from the two models (elasticities in Tables 6a and 7) are significantly different, we use the Hotelling 2-sample T^2 test (Morrisson 1990, Gupta et al. 1996). For this purpose, we first generate 500 draws from the asymptotic distribution of the parameters of both the models. For each draw we computed the 49X1 mean demand elasticity vector (for the seven brands, across store-weeks) for both models. Let \hat{W}_Q denote the empirical variance-covariance matrix of the 500 bootstrapped mean demand elasticities for the proposed model; \hat{W}_B denote the empirical variance-covariance matrix

of the 500 bootstrapped mean demand elasticities for the logit model; and \hat{W}_{QB} denote the pooled variance-covariance matrix of \hat{W}_Q and \hat{W}_B . Let n_1 and n_2 represent the number of observations of the bootstrapped elasticities (in this case 500), and let $\hat{\mu}_Q$ and $\hat{\mu}_B$ be the mean demand elasticity vector computed across the 500 bootstrapped replications for the two models. The null hypothesis of no difference between the elasticities from the two models is tested by the following T^2 -statistic:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\hat{\mu}_Q - \hat{\mu}_B)' \hat{W}_{QB} (\hat{\mu}_Q - \hat{\mu}_B).$$

The quantity $F = ((n_1 + n_2 - p - 1)/(n_1 + n_2 - 2)p)T^2$ has an F distribution with degrees of freedom p and $n_1 + n_2 - p - 1$, where p is the order of the elasticity vectors (here $p = 49$). The computed value of the statistic is 758.52; the corresponding critical value of $F(49, 950)$ at the 0.01 significance level is 1.55, showing that the null hypothesis of equal elasticities is strongly rejected.

5.2.1. Primary Versus Secondary Demand Decompositions. An attractive feature of our specification is the ability to decompose the price elasticity of demand and predicted sales into primary (purchase incidence/quantity) and secondary (brand-switching) demand. In Table 8 we provide both an elasticity-based decomposition (Bell et al. 1999) and a unit sales-based decomposition (van Heerde et al. 2003). Under the elasticity-based decomposition, the primary demand effect is computed as the purchase incidence and purchase quantity elasticity as a percentage of the total demand elasticity, averaged across all store weeks. Under the unit sales decomposition, the primary demand effect is computed as the change in category sales as a percentage of the change in own-brand sales in response to the price change, averaged across store-weeks.¹² Our primary demand component accounts for about 35% of total elasticity and 92% of total unit sales. These values lie in the range reported by Bell et al. 1999 (Table 6) for the elasticity-based decomposition, and by van Heerde et al. 2003 (Table 4) for the unit sales-based decomposition. Consistent with van Heerde et al. (2003), we find that the brand-switching component of unit sales is considerably less than the primary demand (purchase incidence/quantity) component. The comparability of

¹² The expressions for the primary (PD) and secondary (SD) *unit sales* effects as defined in Equations 8 and 9, Van Heerde et al. (2003), in the context of the proposed model are:

$$\begin{aligned} PD_{jst} &= \int \left(1 + \alpha_s \left[1 - \sum_{k=1}^J \frac{p_j}{p_k} \left\{ \Pr(C_{ikst} = 1 | I_{ist} = 1) + \frac{\Pr(C_{ikst} = 1, I_{ist} = 1)}{\ln[\Pr(I_{ist} = 0)]} \right\} \right] \right) \\ &\quad / \left(1 + \alpha_s \left[1 - \Pr(C_{ijst} = j | I_{ist} = 1) - \frac{\Pr(C_{ijst} = 1, I_{ist} = 1)}{\ln[\Pr(I_{ist} = 0)]} \right] \right) d\Omega_i \\ SD_{jst} &= 1 - PD_{jst}. \end{aligned}$$

¹¹ The model structure implies that the expected quantity elasticity is: $\epsilon q_{pk} = -1 + \int \alpha_s [\Pr(C_{ijst} = 1, I_{ist} = 1) / \ln\{\Pr(I_{ist} = 0)\}] * [1 + \ln\{\Pr(I_{ist} = 0)\} \Pr(I_{ist} = 0) / \Pr(I_{ist} = 1)] \phi(\Lambda) \partial \Lambda$. Therefore, low unconditional purchase probabilities, $\Pr(C_{ijst} = 1, I_{ist} = 1)$, imply quantity elasticities close to 1.

Table 7 Expected Price-Elasticities of Demand (Across Stores and Weeks) for the Single-Unit (Logit) Model

Effect of change in price of:	On the market share of:							No purchase probability
	MM64	MM96	Dom64	TRPrm64	TRSB64	TRPrm96	FLR	
MM64	−3.0405*	0.1230*	0.1850	0.2800*	0.2680*	0.3400*	0.0057	0.0322
MM96	0.0237	−3.1162*	0.0143	0.0253	0.0251*	0.0373*	0.0019	0.0071
Dom64	0.2120*	0.0909*	−3.0413*	0.1800*	0.2230*	0.2250*	0.0110	0.0421
TRPrm64	0.2580*	0.1260*	0.1440*	−2.7391*	0.3110*	0.3000*	0.0050	0.0272
TRSB64	0.2330*	0.1190*	0.1710*	0.2940*	−3.0330*	0.3190*	0.0054	0.0297
TRPrm96	0.1190*	0.0664*	0.0644*	0.1070*	0.1230*	−2.7334*	0.0003	0.0072
FLR	0.0011	0.0021	0.0018	0.0009	0.0010	0.0002	−3.2501*	0.0092

* Statistically significantly different from zero at the 10% level.

Table 8 Elasticity and Unit-Sales Based Decomposition for the Proposed Model

Brand	Size (oz.)	Elasticity based decomposition ¹		Unit Sales based decomposition ²	
		Primary (%)	Secondary (%)	Primary (%)	Secondary (%)
Minute Maid	64	38.53	61.47	84.34	15.66
Minute Maid	96	32.94	67.06	96.70	3.30
Dominick's	64	36.86	63.14	91.50	8.50
Tropicana Prm	64	37.22	62.78	97.29	2.71
Tropicana SB	64	32.66	67.34	95.56	4.44
Tropicana Prm	96	33.42	66.58	82.14	17.86
Florida	96	31.52	68.48	98.68	1.32
Average for category		34.73	65.27	92.31	7.69

Notes. Following Bell et al. (1999), the primary demand elasticity is the purchase incidence + quantity elasticity, and the secondary demand elasticity is the conditional brand-choice elasticity.

¹ Primary and secondary demand elasticities as percentages of the total demand elasticity (as reported in Bell et al. 1999, Table 5).

² Primary and secondary unit sales effects as percentages of the total unit sales effect (as defined in Equations 8 and 9, van Heerde et al. 2003).

our results, based on aggregate data, with previous studies, based on household data, serve as a source of validation for our approach.

5.3. Category Purchase Probabilities

We now compare the category purchase probabilities predicted by the proposed model and the logit model. Since the latter only allows one unit to be purchased per customer, we expect it would overpredict purchase incidence (the proportion of store trips resulting in a purchase). Indeed, the mean category purchase probability across all store-weeks estimated at the final parameter values for the single-unit model is 0.0691, about twice the corresponding value of 0.0379 for the proposed model. We use the 2-sample T^2 test described in the previous section to test if the predicted mean category purchase probabilities across the 30 stores in our sample are significantly different between the two models. The computed value of the statistic using the 500 bootstrapped values was 2031.44; the corresponding critical value of $F(30, 969)$ at the 0.01 significance level is 1.72, showing that the null hypothesis of equal category purchase probabilities is strongly rejected.

In Figure 2, we plot the weekly predicted category purchase probabilities for the two models in our sample Store 2. In addition to predicting higher purchase incidence overall, the logit also predicts occasional large spikes in incidence which seem unrealistic. The

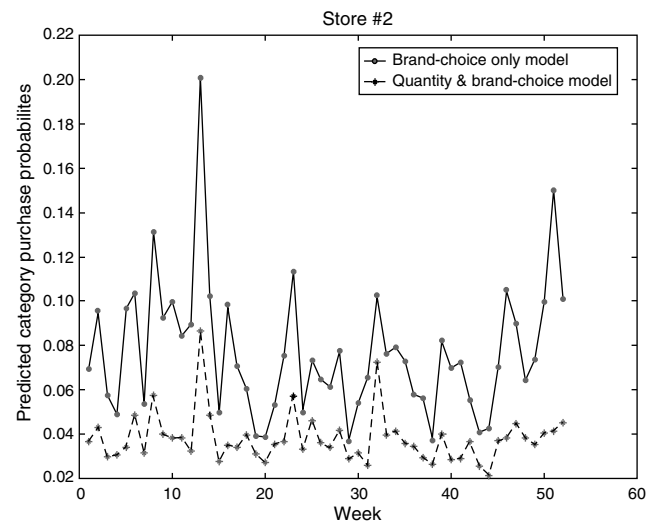
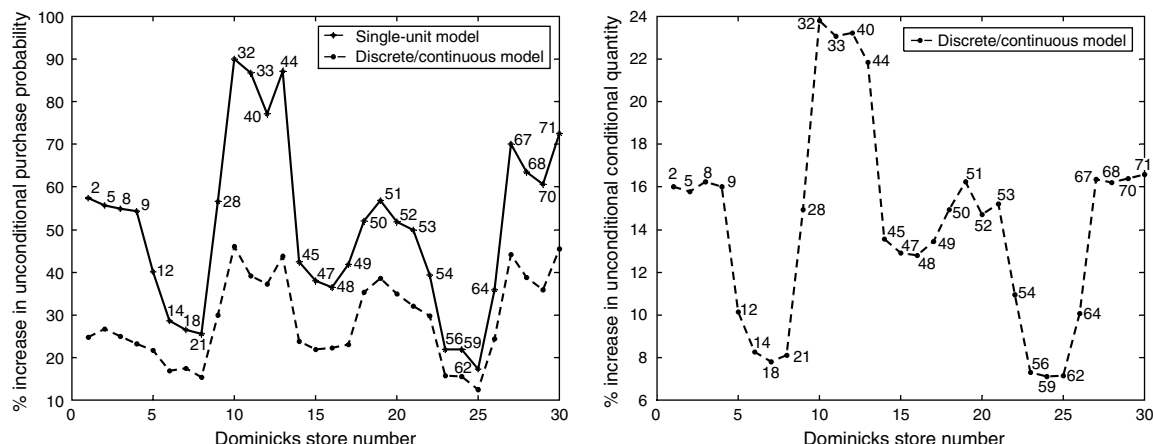
Figure 2 Plot of Predicted Category Purchase Probability Across Weeks for Store 2

Figure 3 Plot of Predicted Percent Increase in Unconditional Choice Probabilities and Expected Conditional Quantities Demanded Across All Stores in Response to a 30-Cent Price Cut of the Store Brand



proposed model smoothes these spikes by allowing for changes in quantity choices.

5.4. Measuring the Impact of a Price Cut

We now illustrate the substantive differences of these two models by simulating the impact of a hypothetical retail price cut of 30 cents for the store brand. Such a price cut might be used as an investment to generate new trials of a product by generating incremental sales or by stealing share from national brands. Figure 3 shows the percentage increase in the unconditional brand-choice probability (left panel) and the conditional quantity (right panel) of the store brand across the 30 stores in response to the price cut. As expected, the logit predicts a much larger degree of category expansion and brand switching. Looking at the right panel, we see that the percentage increase in average conditional quantities demanded in response to this price cut could be as high as 24% (Store 32), indicating that quantity effects are not negligible. By constraining these to be zero a priori, the logit could potentially present the retailer an overly optimistic view of his ability to generate new trials.

6. Conclusions

This paper makes three contributions. First, we present a methodology for estimating the aggregate demand system corresponding to the models of consumer choice of Chiang (1991) and Chintagunta's (1993). This approach provides a parsimonious representation of demand, while retaining the link to consumer theory. Moreover, the approach allows one to recover the structural parameters of the model using aggregate store-level data. Second, our specification allows us to control for the role of heterogeneity in consumer tastes as well as the potential endogeneity of prices, both of which could bias our parameter estimates when not properly accounted for. We propose a modified inversion procedure similar in spirit

to Berry et al. (1995) to control for the endogeneity. Third, our empirical results using weekly store-level data for the refrigerated juice category obtained from the Dominick's Finer Foods database indicate that the model provides reasonable estimates for the average conditional purchase quantity for this market and category. Further, the proposed specification outperforms the traditional logit model along several dimensions.

We expect the logit model would be a reasonable demand specification for many categories in which single-unit purchase behavior is common. However, our empirical results and simulation experiments indicate the need for caution in imposing single-unit purchase behavior in categories for which the assumption is inappropriate. In particular, misusing this assumption leads to erroneous results for price elasticities and category expansion effects. In contrast, our proposed approach serves as a more robust demand specification for packaged-good categories as it can accommodate substitution patterns across a wider scope of quantity choice behaviors.

Notwithstanding the paper's contributions, future applications of our proposed methodology depend on the availability of appropriate instruments. Given that cost drivers and wholesale cost data are increasingly becoming available, this should not be too much of a problem in the future. An interesting methodological extension would be to account explicitly for the discrete nature of purchase quantities. A methodology similar to Arora et al. (1998) could be the starting point for such an analysis. Using the structural derivation of our model, one could measure consumer welfare while accounting for differing welfare gains from primary versus secondary demand considerations.

A limitation of the current analysis and, more generally, comparable models capturing quantity choices is the simplistic treatment of consumer shopping dynamics. Currently, our model does not distinguish between increased consumption versus stock-piling

during weeks with above-average quantity purchases. Accounting for consumer stock-piling and the resulting consumer shopping dynamics is computationally beyond the scope of the current paper. However, important advances have been made in recent work using household data (e.g., Erdem et al. 2003). Extending such dynamic analysis to aggregate settings would be a challenging, but important contribution to this literature.

Acknowledgments

The authors thank Jeongwen Chiang for insightful comments and suggestions that improved an earlier draft of this paper. They also thank Dennis Carlton, Phil Haile, Nitin Mehta, Amil Petrin, and seminar participants at the University of Chicago and the Marketing Science Conference, Edmonton, Alberta, Canada 2002, for useful comments and discussions. The authors also thank Peter Rossi for making the Dominick's data available. The first author gratefully acknowledges research support from the Sanford J. Grossman Fellowship in Honor of Arnold Zellner; any opinions herein are the author's and not necessarily those of Sanford J. Grossman or Arnold Zellner. The second and third authors acknowledge financial support from the James M. Kilts Center for Marketing for Research. All errors remain the authors'.

Appendix

A. Inversion of the Demand Function

In this section of the appendix, we show that our proposed procedure (Equation (12)) for inverting the expected demand function (Equation (11)) to recover δ is a contraction mapping. That is, iteratively solving (12) converges to a vector δ that uniquely reconciles \tilde{Q}_{jst} in Equation (11) with the average quantity per customer in the data, q_{jst} .

The expected demand function (Equations (10) and (11)) implied by the model is:

$$\begin{aligned}\tilde{Q}_{jst}(\delta; \Theta) &= \int -\frac{y_{st}}{\alpha_s p_{jst}} \Pr(D_{ijst} = 1 | I_{ist} = 1) \ln[P(I_{ist} = 0)] \phi(\Lambda) \partial \Lambda \\ &= \int \frac{y_{st}}{\alpha_s p_{jst}} \frac{e^{V_{ijst}}}{\sum_{k=1}^J e^{V_{ikst}}} \ln \left[1 + \sum_{k=1}^J e^{V_{ikst}} \right] \phi(\Lambda) \partial \Lambda \\ &= \int \frac{y_{st}}{\alpha_s p_{jst}} \frac{e^{\delta_{jst} + \Omega_{jst}}}{\sum_{k=1}^J e^{\delta_{kst} + \Omega_{ikst}}} \ln \left[1 + \sum_{k=1}^J e^{\delta_{kst} + \Omega_{ikst}} \right] dF(\Omega).\end{aligned}\quad (A0)$$

The iterative function $g(\cdot): \mathbb{R}^J \rightarrow \mathbb{R}^J$ (Equation (12)) is defined as:

$$g(\delta) = \delta + \ln(q) - \ln[\tilde{Q}(\delta)] \quad (A1)$$

We show that $g(\cdot)$ is a contraction mapping by proving that it satisfies the conditions described in Appendix 1 of Berry et al. (1995). The subscripts s for "store" and t for "week" are dropped for clarity. The main conditions to prove are:

- (a) $g(\cdot)$ is continuous in δ ;
- (b) $\partial g_j(\delta) / \partial \delta_j \geq 0 \quad \forall j$; and,
- (c) $\sum_{r=1}^J \partial g_i(\delta) / \partial \delta_r < 1$.

The function is continuous by construction. To prove (b), we first show that $\partial[g_j(\delta)] / \partial \delta_j \geq 0$, and then that $\partial[g_j(\delta)] / \partial \delta_r \geq 0, \forall r \neq j$.

(b.1) To show that $\partial g_j(\delta) / \partial \delta_j \geq 0$, note that

$$\begin{aligned}\frac{\partial}{\partial \delta_j} \left(\frac{e^{\delta_j + \Omega_{ij}}}{\sum_{k=1}^J e^{\delta_k + \Omega_{ik}}} \right) \\ = \Pr(C_i = j | I_i = 1) * [1 - \Pr(C_i = j | I_i = 1)],\end{aligned}$$

and that

$$\frac{\partial}{\partial \delta_j} \left(\ln \left[1 + \sum_{k=1}^J e^{\delta_k + \Omega_{ik}} \right] \right) = \Pr(C_i = j, I_i = 1).$$

Also note that at a given guess of the parameter vector, α is known and is fixed. Hence,

$$\begin{aligned}\partial[g_j(\delta)] / \partial \delta_j \\ = 1 - \frac{1}{\tilde{Q}_j(\cdot)} \int \frac{y}{\alpha p_j} \Pr(C_i = j | I_i = 1) [\Pr(C_i = j, I_i = 1) \\ - \ln\{\Pr(I_i = 0)\} \{1 - \Pr(C_i = j | I_i = 1)\}] dF(\Omega).\end{aligned}\quad (A2)$$

Comparing the numerator of the second term above to line 1 of Equation (A0), we can see that for $\partial[g_j(\delta)] / \partial \delta_j \geq 0$, it is equivalent to prove that:

$$\begin{aligned}[\Pr(C_i = j, I_i = 1) - \ln\{\Pr(I_i = 0)\} \{1 - \Pr(C_i = j | I_i = 1)\}] \\ \leq -\ln\{\Pr(I_i = 0)\},\end{aligned}$$

which is equivalent to showing that

$$\Pr(C_i = j | I_i = 1) [1 - \Pr(I_i = 0) + \ln\{\Pr(I_i = 0)\}] \leq 0.$$

This holds since $\Pr(I_i = 0) \in [0, 1]$.

(b.2) To show that $\partial g_j(\delta) / \partial \delta_r \geq 0, r \neq j$, note first that

$$\begin{aligned}\frac{\partial}{\partial \delta_r} \left(\frac{e^{\delta_j + \Omega_{ij}}}{\sum_{k=1}^J e^{\delta_k + \Omega_{ik}}} \right) \\ = -\Pr(C_i = j | I_i = 1) \Pr(C_i = r | I_i = 1) \Rightarrow \frac{\partial g_j(\delta)}{\partial \delta_r} \\ = -\frac{1}{\tilde{Q}_j(\cdot)} \int \frac{y}{\alpha p_j} \Pr(C_i = j | I_i = 1) \Pr(C_i = r | I_i = 1) \\ \cdot [\Pr(I_i = 1) + \ln\{\Pr(I_i = 0)\}] dF(\Omega) \\ = \frac{1}{\tilde{Q}_j(\cdot)} \int \frac{y}{\alpha p_j} \Pr(C_i = j | I_i = 1) \Pr(C_i = r | I_i = 1) \\ \cdot \underbrace{(-\ln\{[1 - \Pr(I_i = 1)] e^{\Pr(I_i = 1)}\})}_{\geq 0} dF(\Omega) \\ \geq 0.\end{aligned}\quad (A3)$$

(c) To show that the sum of the derivatives is less than 1, note that:

$$\sum_{r=1}^J \frac{\partial g_i(\delta)}{\partial \delta_r} = 1 - \frac{1}{\tilde{Q}_i(\cdot)} \int \frac{y}{\alpha p_i} \Pr(C_i = i | I_i = 1) f_i dF(\Omega), \quad (A4)$$

where,

$$\begin{aligned}f_i &= \Pr(C_i = j, I_i = 1) - \ln\{\Pr(I_i = 0)\} \{1 - \Pr(C_i = j | I_i = 1)\} \\ &\quad + \{\Pr(I_i = 1) + \ln\{\Pr(I_i = 0)\}\} \sum_{r \neq j} \Pr(C_i = r | I_i = 1).\end{aligned}$$

Comparing the numerator of the second term in (A4) to line 1 of Equation (A0), we can see that for $\sum_{r=1}^J \partial g_i(\delta) / \partial \delta_r < 1$, it is equivalent to prove that: $f_i < -\ln\{\Pr(I_i = 0)\}$. This reduces to showing that

$$\Pr(I_i = 1) * [\Pr(I_i = 1) + \ln\{1 - \Pr(I_i = 1)\}] < 0.$$

This holds since $\Pr(I_i = 1) \in [0, 1]$.

Hence, $g(\cdot)$ is a contraction mapping, and therefore, iteratively solving (A1) will converge to a unique vector δ .

B. Method of Simulated Moments Procedure

Recall from §3 that we construct moment conditions based on the mean-independence assumption $E[\xi_{jst} Z_{jst} | Z_{jst}] = 0$, where Z_{jst} contains both the observed product characteristics and additional exogenous instruments. Since the evaluation of $\xi(\Theta)$ requires simulation of integrals, estimation is carried out using a method of simulated moments procedure (Pakes and Pollard 1989). We define the sample analogue of the moment vector, $M(\Theta) = Z' \xi(\Theta)$. Estimation involves searching for parameter values, Θ^{GMM} , that set $M(\Theta)$ as close as possible to zero. We estimate all parameters by minimizing the GMM objective function, $\Omega(\Theta) = M(\Theta)' W M(\Theta)$. For our application, we use the optimal weight matrix, $W = (E[M(\Theta) M(\Theta)'])^{-1}$ (Hansen 1982).

C. Simulations

We conduct several simulations to assess the robustness of our proposed estimation procedure to various forms of consumer behavior. In each case, we use 30 replications and report the mean findings across the replications. First, we show that the proposed model is capable of recovering the parameters when it is the true model. Second, we show the model provides reasonable price elasticities when the “true” model is the logit. Third, in contrast with the second simulation, we show that the logit model does not perform well when the “true” data are generated from the proposed model. Fourth, we show that the proposed specification performs reasonably well when the data are generated from an alternative (noneconomic) model of consumer behavior. Finally, we show that the proposed model is robust to indivisible consumer quantity choices at the individual level. A more detailed account of these simulations and the results (along with additional simulation cases) is available online at the journal website.

(1) Recovery of Parameters

We begin by generating (simulated) data from the proposed model to verify that our MSM procedure is capable of recovering the true parameter values. We consider both the cases with and without unobserved taste heterogeneity for brands (i.e., random coefficients). For the latter, we also consider the case that the unobserved brand characteristics, ξ_{jst} , are uncorrelated with prices.

For simplicity, we assume there are two alternatives in the choice set and that prices and promotions are the only causal variables other than intrinsic brand tastes. Actual scanner data from a two-brand product category are used for prices and promotions. The data consist of prices and promotions at the chain level for 90 weeks for the oats category, in which Quaker Oats and the Dominick’s store brand

are the only two major brands. Price and promotional variables for the model were created by taking the difference of the variables across the two brands. Average expenditures of consumers for each of the 90 weeks were simulated as *Uniform*(5, 20). Average quantities per consumer for the 90 weeks were generated by integrating over the expected demands of consumers who are allowed to make multiple-unit purchases. We report the results in Table A1 and we find that the proposed model does a good job of recovering all the parameters.

We now allow for unobserved heterogeneity in the intrinsic brand preference parameter.¹³ We consider two cases corresponding to low/high variance in the intrinsic preference heterogeneity. The results are reported in Table A2 and correspond to the means, standard deviations, and mean absolute percentage deviations (MAPDs) of the recovered parameter values for the two cases. The results reveal that for the range of parameter values considered, the proposed model does a good job of recovering the intrinsic preference and the price and promotion sensitivity parameters. The variation in the standard deviation in the intrinsic preference heterogeneity across replications is high, but is comparable to those of the single-unit logit (see Chintagunta 2003). We conclude that our proposed model can recover the true model parameters from the aggregate data in the presence of multiple-unit purchases.

(2) Performance of Proposed Model When Data Are Generated from Logit

Here we generate data from the standard logit demand system and to see how well the proposed discrete/continuous model performs (i.e., the true data consist of single-unit purchases but the econometric model allows for multi-unit purchases). One must keep in mind that not all of the parameters are directly comparable across the two models (i.e., logit and discrete/continuous). Some of the structural interpretations of parameters differ as do the role these parameters play in the statistical models. For instance, the price parameter represents the marginal utility of income in the logit. Statistically, it captures the price sensitivity of choices. In the discrete/continuous model, the price parameter, α_s , is a scale parameter from the quality function. Statistically, it captures both the price sensitivities of choices and quantities. We will report parameter estimates from both a logit and a discrete/continuous specification below. But, to make a more sensible comparison, we will focus our attention on comparing elasticities. We compare category demand expansion/brand switching effects implied by the proposed discrete/continuous model when it is estimated on data generated by logit demand.

We consider a 3-alternative case with 2 brands and an outside good. 190 weeks of prices for the two brands are generated as *Uniform*(2, 5); demand shocks simulated as *normal*(0, 0.1), and expenditures (y) fixed at 5. Using this “data,” market shares corresponding to various sets of brand intercepts and price sensitivity parameters are simulated from a logit model. The proposed discrete/continuous model is then estimated. The implied category purchase

¹³ The number of draws for simulating the integral was fixed at 100 and the variance of ξ is set to 1 for all simulations.

Table A1 Monte Carlo Results with No Heterogeneity

	True and estimated parameter values for the discrete/continuous model											
	True	Model	True	Model	True	Model	True	Model	True	Model	True	Model
Intrinsic preference	−4.50	−4.466	−4.50	−4.458	−4.50	−4.473	−4.50	−4.480	−4.50	−4.480	−4.50	−4.483
Price sensitivity	1.50	1.545	2.00	2.050	2.50	2.539	3.00	3.032	3.50	3.547	4.00	4.037
Promotion sensitivity	0.50	0.418	0.50	0.413	0.50	0.418	0.50	0.402	0.50	0.425	0.50	0.428

Table A2 Monte Carlo Results with Heterogeneity in Intrinsic Preferences

	Low std. deviation in intrinsic preference				High std. deviation in intrinsic preference			
	True	Mean	Std. deviation	MAPD	True	Mean	Std. deviation	MAPD
Intrinsic preference	−6.000	−6.536	1.200	24.637	−6.000	−5.665	1.297	28.650
Price sensitivity	2.000	1.897	0.609	14.430	2.000	1.637	0.626	19.274
Promotion sensitivity	0.500	0.652	0.382	18.508	0.500	0.563	0.671	21.641
Std. deviation in intrinsic preference	0.500	0.303	0.811	67.498	1.000	0.305	1.298	84.750

Table A3 Performance of Proposed Discrete/Continuous Model When Data Are Generated from Logit

	Due to a 1% ↑ in the price of:	True		True		True		True		True	
		True	Model	True	Model	True	Model	True	Model	True	Model
Brand 1 intercept		−2.000	−3.646	−2.000	−3.227	−2.000	−2.926	−2.000	−2.699	−2.000	−2.515
Brand 2 intercept		−4.000	−5.650	−4.000	−5.230	−4.000	−4.928	−4.000	−4.700	−4.000	−4.516
Log(Price)		−2.000	−1.779	−2.500	−2.283	−3.000	−2.814	−3.500	−3.396	−4.000	−3.835
% change in conditional choice probability of:											
Brand 1	Brand 1	−0.279	−0.240	−0.360	−0.319	−0.533	−0.484	−0.601	−0.573	−0.807	−0.754
Brand 2	Brand 2	−1.700	−1.521	−2.111	−1.939	−2.430	−2.296	−2.850	−2.775	−3.135	−3.027
Brand 2	Brand 1	1.725	1.542	2.150	1.971	2.481	2.342	2.922	2.843	3.221	3.106
Brand 1	Brand 2	0.276	0.238	0.355	0.315	0.527	0.478	0.593	0.566	0.798	0.745
% change in category purchase probability:											
	Brand 1	−1.670	−1.512	−2.084	−1.926	−2.407	−2.281	−2.826	−2.757	−3.109	−3.004
	Brand 2	−0.271	−0.236	−0.351	−0.313	−0.520	−0.473	−0.585	−0.559	−0.785	−0.735
Brand 1 intercept		−2.000	−2.917	−2.000	−2.919	−2.000	−2.935	−2.000	−2.831	−2.000	−2.947
Brand 2 intercept		−2.000	−2.902	−2.500	−3.439	−3.000	−3.931	−3.500	−4.346	−4.000	−4.947
Log(Price)		−3.000	−2.756	−3.000	−2.898	−3.000	−2.851	−3.000	−2.742	−3.000	−2.729
% change in conditional choice probability of:											
Brand 1	Brand 1	−1.482	−1.364	−1.151	−1.109	−0.815	−0.771	−0.673	−0.596	−0.510	−0.443
Brand 2	Brand 2	−1.494	−1.372	−1.823	−1.765	−2.154	−2.053	−2.293	−2.117	−2.453	−2.253
Brand 2	Brand 1	1.503	1.379	1.845	1.785	2.190	2.085	2.337	2.154	2.505	2.297
Brand 1	Brand 2	1.491	1.370	1.152	1.109	0.811	0.766	0.667	0.590	0.503	0.438
% change in category purchase probability:											
Brand 1	−1.463	−1.350	−1.793	−1.744	−2.125	−2.033	−2.267	−2.099	−2.429	−2.238	
	Brand 2	−1.450	−1.342	−1.126	−1.090	−0.797	−0.756	−0.657	−0.583	−0.497	−0.434

Notes. Data for this table: 190 weeks of data generated from logit demand model with Brand 1 prices = uniform(2, 5), Brand 2 prices = uniform(2, 5), $\gamma = 5$, $\xi = \text{iid } N(0, 0.1)$. “True” = Logit; “Model” = Discrete/continuous.

incidence, and conditional (on category purchase) brand-choice elasticities for both models are computed by taking the average of the percent change in these values due a 1% change in price, across all observations. The results are presented in the two panels in Table A3. Since parameters are not directly comparable across the logit and discrete/continuous models, we focus instead on the estimated elasticities. Referring to Table A3, we find that the discrete/continuous model is reasonably able to recover the incidence and brand-choice elasticities implied by the “true” logit model that generated the data.

(3) Performance of Logit When Data Are Generated from Proposed Model

We now consider the reverse case to the second simulation experiment. We simulate data from the discrete/continuous demand model and then estimate the parameters using a logit model. Analogous to study (b), we compare category demand expansion/brand switching effects. The results are presented in Table A4. We find that when the true underlying model is discrete/continuous, the logit does not seem to provide reasonable elasticities. In particular, when conditional quantities at the individual level range from 2 to

Table A4 Performance of Logit When Data Are Generated from Proposed Model with High Conditional Quantities

Due to a 1% increase in price of:		True	Model	True	Model	True	Model	True	Model	True	Model
Brand 1 intercept		−1.000	2.418	−1.000	1.707	−1.000	1.298	−1.000	1.098	−1.000	0.872
Brand 2 intercept		−1.000	2.421	−1.000	1.689	−1.000	1.305	−1.000	1.069	−1.000	0.868
Log(Price)		−2.000	−3.610	−2.500	−3.778	−3.000	−4.152	−3.500	−4.586	−4.000	−5.020
Mean conditional quantity		4.047		3.102		2.590		2.216		1.919	
% change in conditional choice probability of:											
Brand 1	Brand 1	−0.979	−1.743	−1.284	−1.962	−1.474	−2.021	−1.771	−2.332	−2.049	−2.571
Brand 2	Brand 2	−1.009	−1.832	−1.199	−1.778	−1.502	−2.083	−1.697	−2.195	−1.908	−2.377
Brand 2	Brand 1	1.012	1.850	1.203	1.794	1.512	2.112	1.711	2.228	1.929	2.420
Brand 1	Brand 2	0.981	1.759	1.290	1.984	1.482	2.047	1.787	2.371	2.075	2.623
% change in category purchase probability:											
Brand 1	−0.931	−1.296	−1.138	−1.512	−1.452	−1.927	−1.647	−2.066	−1.863	−2.288	
	Brand 2	−0.902	−1.233	−1.220	−1.668	−1.421	−1.848	−1.722	−2.205	−2.002	−2.474

Notes. 190 weeks of data generated from logit demand model with Brand 1 prices = uniform(2, 5), Brand 2 prices = uniform(2, 5), $\gamma = 25$, $\xi = \text{iid } N(0, 0.1)$. “True” = Discrete/continuous; “Model” = Logit.

4 units, the logit overstates aggregate conditional brand-choice elasticities by about 44%, and aggregate category purchase elasticities by about 31% on average. The overstating of the brand-choice and category expansion effects is consistent with the empirical results in the paper (see §§5.2 and 5.3).

(4) Divisibility of Quantities

Finally, we generate data from an underlying model in which quantity choices are indivisible (i.e., integer quantity choices) to explore the robustness of the divisibility assumption implicit in our proposed specification. We generate data from the model of Arora et al. (1998) (henceforth AAG). For simplicity, here we consider the one alternative case in the AAG framework.

The aggregate demand corresponding to AAG resembles the proposed model except for the expression for expected conditional quantities. Specifically, the conditional quantities are (Equation 11, p. 33 in AAG):

$$p_0(q) = \Pr(Q_j = q) = F[(q + 0.5)p_j/\gamma] - F[(q - 0.5)p_j/\gamma],$$

where $F(\cdot)$ is the cdf of an extreme value distribution with location parameter $\alpha/\gamma - \ln(p) + \mu \ln[1/\Pr(j)]$ and scale parameter μ . The corresponding expected conditional quantity to this model is:

$$E[Q]_{\text{discrete}} = \sum_{q=1}^{\infty} [q * p_0(q)].$$

In contrast, our proposed specification has expected conditional quantity:

$$E[Q]_{\text{continuous}} = \frac{\gamma}{p_j} [\alpha/\gamma - \ln(p_j) + \mu(\Gamma + \ln\{1/\Pr(j)\})],$$

where Γ is Euler’s constant. In the simulation, we compare $E[Q]_{\text{discrete}}$ and $E[Q]_{\text{continuous}}$. We first generate N (= 100) prices from uniform(2, 4), and compute $E[Q]_{\text{discrete}}$ and $E[Q]_{\text{continuous}}$ for various values of α/γ and γ . For all computations, $\mu = 1$, and “ ∞ ” = 50 (in the summation on $E[Q]_{\text{discrete}}$). The results are given in Table A5. We see that

Table A5 Comparison of Expected Conditional Quantity Under Discrete and Continuous Cases

$\alpha/\gamma = 3.0$			$\gamma = 1.0$		
γ	$E[Q]_{\text{continuous}}$	$E[Q]_{\text{discrete}}$	α/γ	$E[Q]_{\text{continuous}}$	$E[Q]_{\text{discrete}}$
1.0	0.9677	0.9833	3.0	0.8960	0.9150
1.1	1.0644	1.0807	3.1	0.9257	0.9492
1.2	1.1612	1.1767	3.2	0.9558	0.9831
1.3	1.2580	1.2720	3.3	0.9862	1.0166
1.4	1.3547	1.3671	3.4	1.0169	1.0496
1.5	1.4515	1.4623	3.5	1.0479	1.0820
1.6	1.5483	1.5577	3.6	1.0791	1.1137
1.7	1.6450	1.6533	3.7	1.1106	1.1447
1.8	1.7418	1.7491	3.8	1.1422	1.1750
1.9	1.8386	1.8451	3.9	1.1741	1.2047
2.0	1.9353	1.9413	4.0	1.2062	1.2339

the difference in expected conditional quantity between the two cases is quite small. And since it is the expected quantity that impacts the aggregate demand functions, the difference between the two cases is likely to be very small.

References

- Allenby, Greg M., Thomas S. Shively, Sha Yang, Mark J. Garratt. 2004. A choice model for packaged goods: dealing with discrete quantities and quantity discounts. *Marketing Sci.* 23(1) 95–108.
- Allenby, G., Peter E. Rossi. 1991. Quality perceptions and asymmetric switching between brands. *Marketing Sci.* 10 185–205.
- Arora, N., G. Allenby, J. L. Ginter. 1998. A hierarchical Bayes model of primary and secondary demand. *Marketing Sci.* 17 29–44.
- Bell, David R., Jeongwen Chiang, V. Padmanabhan. 1999. The decomposition of promotional response: an empirical generalization. *Marketing Sci.* 18(4) 504–526.
- Berry, S. 1994. Estimating discrete choice models of product differentiation. *Rand J. Econom.* 25(2) 242–262.
- Berry, S. T., J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* 60(4) 841–890.
- Besanko, D., S. Gupta, D. Jain. 1998. Logit demand estimation under competitive pricing behavior: an equilibrium framework. *Management Sci.* 44(11) 1533–1547.

- Bhat, C. R. 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Res.* **35B** 677–693.
- Blattberg, R., K. Wisniewski. 1989. Price-induced patterns of competition. *Marketing Sci.* **8** 291–310.
- Chiang, J. 1991. The simultaneous approach to the whether, what, and how much to buy questions. *Marketing Sci.* **10** 297–315.
- Chintagunta, P. K. 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Sci.* **12** 184–208.
- Chintagunta, P. K. 2003. A flexible aggregate logit demand model. Working paper, Graduate School of Business, University of Chicago, Chicago, IL.
- Chintagunta, P. K., Jean-Pierre Dubé, V. Singh. 2003. Balancing profitability and customer value: an application to zone-pricing by a supermarket chain. *Quant. Marketing Econom.* **1**(1) 111–117.
- Conley, T. G. 1999. GMM estimation with cross sectional dependence. *J. Econom.* **92** 1–45.
- Dillon, W. R., S. Gupta. 1996. A segment-level model of category volume and brand choice. *Marketing Sci.* **15**(1) 38–59.
- Dubé, J. P. 2004. Multiple discreteness and product differentiation: demand for carbonated soft drinks. *Marketing Sci.* **23**(1) 26–81.
- Erdem, T., S. Imai, M. Keane. 2003. A model of consumer brand and quantity choice dynamics under price uncertainty. *Quant. Marketing Econom.* **1**(1) 5–64.
- Gupta, S., P. Chintagunta, A. Kaul, D. Wittink. 1996. Do household scanner data provide representative inferences from brand choices: a comparison with store data. *J. Marketing Res.* **3** 383–393.
- Hanemann, W. M. 1984. Discrete/continuous models of consumer demand. *Econometrica* **52** 541–561.
- Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* **50** 1029–1054.
- Hendel, I., Aviv Nevo. 2003. Sales and consumer inventory. Working paper, Department of Economics, Northwestern University, Evanston, IL.
- Hoch, S. J., B.-D. Kim, A. L. Montgomery, P. E. Rossi. 1995. Determinants of store-level price elasticity. *J. Marketing Res.* **32** 17–30.
- Kim, J., G. M. Allenby, P. E. Rossi. 2002. Modeling consumer demand for variety. *Marketing Sci.* **21**(3) 229–250.
- Krishnamurthi, L., S. P. Raj. 1988. A model of brand choice and purchase quantity price sensitivities. *Marketing Sci.* **7** 1–20.
- Morrison, Donald. 1996. *Multivariate Statistical Methods*. McGraw-Hill, New York.
- Mehta, N., B. P. S. Murthi, S. Rajiv, K. Srinivasan. 1998. A structural model of quantity and brand choice: estimation issues and managerial implications. Working paper, Rotman School of Management, University of Toronto, Ontario, Canada.
- Pakes, A., D. Pollard. 1989. Simulation and the asymptotics of optimization estimators. *Econometrica* **57**(5) 1027–1057.
- Pollak, Robert A., T. A. Wales. 1992. Demand system specification and estimation. Oxford University Press, New York, 53–66.
- Sudhir, K. 2001. Structural analysis of competitive pricing in the presence of a strategic retailer. *Marketing Sci.* **20**(3) 244–264.
- van Heerde, H., Sachin Gupta, Dick Wittink. 2003. Is 3/4 of the sales promotion bump due to brand switching? No, it is 1/3. *J. Marketing Res.* **40**(4) 481–491.
- van Heerde, Harald J., Peter S. H. Leeflang, Dick R. Wittink. 2004. Decomposing the sales promotion bump with store data. *Marketing Sci.* **23**(3) 317–334.