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

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). 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 stockpile 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.

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 gate demand system. A utility-maximizing consumer following section, we derive the corresponding aggregate demand system. Then, we derive the corresponding aggregate demand system. The random disturbances, \( \varepsilon_{jst} \) and \( \varepsilon_{ist} \), 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 \( \varepsilon_{ist} = (\varepsilon_{j1st}, \ldots, \varepsilon_{jnst}) \sim \text{EV}(0, \mu_j) \) and \( \varepsilon_{ist} \sim \text{EV}(D_i, \tau, \mu) \), where \( D_i \) is a vector of demographic variables specific to store \( s \), and \( \tau \) is a vector of parameters to be estimated. 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 \( \xi_{ist} \). The error component \( \xi \) 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, \( y_{ist} \), and the “quality-weighted” prices of the chosen brand, \( p_{jst}/\psi_{ist} \), and the numeraire good, \( 1/\psi_{ist} \). Using Roy’s identity, the demand function (conditional on category purchase, \( I_{jst} = 1 \), and choice of brand \( j \), \( C_{jst} = 1 \)) corresponding to the HITL is:

\[
\begin{align*}
& x_{jst}(p_{jst}, \psi_{ist}, y_{ist}) = \alpha_1 - \alpha_3 \ln (p_{jst}/\psi_{ist}) + \alpha_3 \ln (1/\psi_{ist}), \\
& = (y_{ist}/p_{jst})[\alpha_1 - \alpha_3 \ln (p_{jst}/\psi_{ist}) + \alpha_3 \ln (1/\psi_{ist})],
\end{align*}
\]

where \( \alpha_1 \) and \( \alpha_3 \) 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

\[
E_j(x_{jst}) = \int_{\text{chosen}} x_{jst}(p_{jst}, \psi_{ist}, y_{ist}) | C_{jst} = 1, I_{jst} = 1 \, dx_{jst} \]

Here \( \alpha_1 \) is a store-specific scale that shifts the perceived quality of brands and the outside good across stores; \( \gamma_{ist} \) is the intrinsic taste for brand \( j \) for a consumer \( i \) in store-week \( st \) and \( d_{ist} \) is a deal variable that indicates an in-store display for brand \( j \) in store-week \( st \) with corresponding consumer sensitivity parameter \( \beta_{ist} \). The random disturbances, \( \varepsilon_{ist} \) and \( \varepsilon_{jst} \), 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 \( \varepsilon_{ist} = (\varepsilon_{j1st}, \ldots, \varepsilon_{jnst}) \sim \text{EV}(0, \mu_j) \) and \( \varepsilon_{ist} \sim \text{EV}(D_i, \tau, \mu) \), where \( D_i \) is a vector of demographic variables specific to store \( s \), and \( \tau \) is a vector of parameters to be estimated. 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 \( \xi_{ist} \). The error component \( \xi \) 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, \( y_{ist} \), and the “quality-weighted” prices of the chosen brand, \( p_{jst}/\psi_{ist} \), and the numeraire good, \( 1/\psi_{ist} \). Using Roy’s identity, the demand function (conditional on category purchase, \( I_{jst} = 1 \), and choice of brand \( j \), \( C_{jst} = 1 \)) corresponding to the HITL is:

\[
\begin{align*}
& x_{jst}(p_{jst}, \psi_{ist}, y_{ist}) = \alpha_1 - \alpha_3 \ln (p_{jst}/\psi_{ist}) + \alpha_3 \ln (1/\psi_{ist}), \\
& = (y_{ist}/p_{jst})[\alpha_1 - \alpha_3 \ln (p_{jst}/\psi_{ist}) + \alpha_3 \ln (1/\psi_{ist})],
\end{align*}
\]

where \( \alpha_1 \) and \( \alpha_3 \) 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

\[
E_j(x_{jst}) = \int_{\text{chosen}} x_{jst}(p_{jst}, \psi_{ist}, y_{ist}) | C_{jst} = 1, I_{jst} = 1 \, dx_{jst} \]

\[= \frac{1}{\alpha_1 + \mu \alpha_3} \left[ \ln (P(I_{jst} = 0)) \right] \]

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 \( f \) for week for the demographic variables.
and corresponding choice probabilities

\[
\Pr(I_{ist} = 1) = \sum_{j=1}^{J} e^{V_{ist}} / \left(1 + \sum_{j=1}^{J} e^{V_{ist}} \right);
\]

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

where \( V_{ist} = \gamma_{i, j, s, t} + \alpha_s / \alpha_3 + \beta_{ist} d_{ist} - \alpha_i \ln(p_{jst}) - D_s' \tau + \xi_{ist} \). (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, \( \psi_{ist} \) (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:

\[
(\gamma_{i, j, s, t}, ..., \gamma_{i, j, s, t})' = (\gamma_1, ..., \gamma_J, ..., \gamma_J)' + L \sigma_{ist}
\]

such that \( \sigma_{ist} \sim MVN(0, I_{JxJ}) \) and \( L' \sigma = \sum_{j=1}^{J} \sigma_{ist} \) and,

\[
\beta_{ist} = \beta + \sigma_s s_{ist}, \quad s_{ist} \sim N(0, 1)
\]

\[
\alpha_s = \alpha + D_s' \theta
\]

This formulation of the demand model differs from previous work in two ways. First, we include the error component \( \xi \) 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 is 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.\(^6\)

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

\[
Q_{jst} = N_{st} \int [Pr(C_{ijst} = 1, I_{ist} = 1)E_s(x_{ist})] \phi(\Lambda) d \Lambda,
\]

where \( \Lambda = (\sigma, s)' \) 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):

\[
\tilde{Q}_{jst} = \int [Pr(C_{ijst} = 1, I_{ist} = 1)E_s(x_{ist})] \phi(\Lambda) d \Lambda
\]

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:

\[
\tilde{Q}_{jst} = \int \Pr(C_{ijst} = 1, I_{ist} = 1) \Phi(\Lambda) d \Lambda
\]

Essentially, (9) is equivalent to imposing \( E_s(x_{ist}) = 1 \) 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, \( \xi_{ist} \).

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

\[
\tilde{Q}_{jst}(\Gamma) = \int \frac{y_{ist}}{\alpha_s p_{jst}} \frac{e^{V_{jst}}}{\sum_{k=1}^{K} e^{V_{kst}}} \ln \left[1 + \sum_{k=1}^{K} e^{V_{kst}}\right] \phi(\Lambda) d \Lambda,
\]

where

\[
V_{jst} = [\gamma_j + \alpha_j / \alpha_3 - \alpha_1 \ln(p_{jst}) + d_{jst} \beta + D_s' \alpha_1 / \alpha_3 - D_s' \tau + \xi_{ist}] \frac{1}{\mu} + [L \sigma_{ist} + d_{jst} \sigma_s s_{ist}] \frac{1}{\mu}
\]

and \( \Gamma \) is a vector containing all the model parameters. In principle, one could estimate (10) using maximum likelihood, where \( \xi \) 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 $\xi$ 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:

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

where $\delta_{ijst} = [(\gamma_j + \alpha_3) - \alpha_2 \ln(p_{ijst}) + d_{ijst} + \Delta_j \theta \alpha_3 - \Delta_j \tau + \xi_{ijst}] / \mu$ is the component of $V_{ijst}$ that is common across consumers in store-week $st$, and $\Omega_{ijst} = (L, \sigma_{ijst} + d_{ijst} \sigma_{ijst}) / \mu$ is the consumer-specific component. Due to linearity, we can write $\delta_{ijst}$ as $\delta_{ijst} = X_{ijst}B - \alpha_2 \ln(p_{ijst}) + \xi_{ijst}$, where $B$ contains an intercept term, $\gamma_j / \mu + \alpha_3 \alpha_3 / \mu_3$; a mean sensitivity to display, $\beta / \mu$, and sensitivities to demographic variables. The scale $\mu$ cannot be recovered separately and is normalized to 1. In addition, we normalize $\alpha_3$ to 1 since it cannot be identified separately from $\alpha_2$. The term $\alpha_2 (= \tilde{\alpha} + \Delta_j \theta)$ is henceforth referred to as the price sensitivity for store $s$. The consumer-specific term, $\Omega_{ijst}$, depends on the remaining parameters, $L$, $\sigma_{ijst}$. Denote the entire set of parameters to be estimated as $\Theta = \{B, \tilde{\alpha}, \theta, L, \sigma_{ijst}\}$.

Analogous to Berry (1994), we invert (11) to recover the vector $\delta_{ijst}(\Theta)$. In the appendix we prove that (11) is indeed an invertible function of $\delta_{ijst}$. Hence, for a given set of model parameters and given values for $\xi_{ijst}$, there exists a unique vector $\delta_{ijst}$ such that (11) holds identically (i.e., $Q_{ijst} = \tilde{Q}_{ijst}(\delta; \Theta)$). To invert $\delta_{ijst}$ 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 $q_{ijst}$. 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 $\Theta$, we iterate on (12) to recover the unique vector $\delta_{ijst}(\Theta)$ that solves (11) across all store-weeks. Intuitively, we calibrate values of $\delta_{ijst}(\Theta)$ that exactly fit the predicted $\tilde{Q}_{ijst}$ to the observed $q_{ijst}$ in the data. To evaluate the multidimensional integrals in (11), we use Monte Carlo simulation. Thus, the inversion procedure described above matches the simulated $\tilde{Q}_{ijst}$ to the observed average $q_{ijst}$ in the data. Simulation was carried out using 100 draws.\(^7\)

Using our values of $\delta_{ijst}(\Theta)$, we construct moment conditions based on $\xi_{ijst} = \delta_{ijst}(\Theta) - X_{ijst}B + \alpha_2 \ln(p_{ijst})$, where we assume $E[\xi_{ijst}X_{ijst} \mid X_{ijst}] = 0$. The endogeneity of prices arises if $E[\xi \ln(p) \mid \ln(p)] \neq 0$. Hence, in addition to $X_{ijst}$, we need additional instruments, $Z_{ijst}$, such that $E[\xi_{ijst}Z_{ijst} \mid Z_{ijst}] = 0$. 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 occa-
Sion (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 cor-
relation 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 whole-
sale prices as additional instruments for shelf prices. 
The motivation for using wholesale prices as instru-
ments comes from the interpretation of $\xi_{ijt}$ as store-
specific aggregate shocks (i.e., aggregate shocks for 
all consumers shopping in a given store). We refer 
the reader to Chintagunta et al. (2003) for a discus-
sion. A limitation of this approach is that it would 
not resolve endogeneity arising from more macroeco-
nomic 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 operational-
ize $y_{st}$ 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 accelera-
tion. 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 promo-
tion 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 accelera-
tion (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

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Product-Level Averages Across Store-Weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand</td>
<td>Size (oz.)</td>
</tr>
<tr>
<td>Minute Maid</td>
<td>64</td>
</tr>
<tr>
<td>Minute Maid</td>
<td>96</td>
</tr>
<tr>
<td>Dominick's</td>
<td>64</td>
</tr>
<tr>
<td>Tropicana Prm</td>
<td>64</td>
</tr>
<tr>
<td>Tropicana SB</td>
<td>64</td>
</tr>
<tr>
<td>Tropicana Prm</td>
<td>96</td>
</tr>
<tr>
<td>Florida</td>
<td>96</td>
</tr>
</tbody>
</table>

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 $\xi$. Previous research has routinely rejected the hypothesis that 
displays correlate with $\xi$ 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.
that allows for heterogeneity and accounts for price endogeneity (Table 2). To allow demand to vary across stores, we include demographic variables as follows:

$$a_s = a + \theta_{ethnic} + \theta_{hval150}$$

$$D_s \tau = \tau_{dvrtn} + \tau_{age60} + \tau_{hlarge}$$

where $a_s$ is the price parameter. An unrestricted variance-covariance matrix for the intercepts, $\Sigma$, 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, $E_x|x_{ist}, p_{ist}, \psi_{ist}, \psi_{ist}, y_{id} \mid C_{ist} = 1, I_{ist} = 1$, 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)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$t$-Stat</th>
<th>$t$-Stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minute Maid 64 oz.</td>
<td>-19.278</td>
<td>8.145</td>
</tr>
<tr>
<td>Minute Maid 96 oz.</td>
<td>-5.676</td>
<td>2.905</td>
</tr>
<tr>
<td>Dominick’s 64 oz.</td>
<td>-14.250</td>
<td>7.474</td>
</tr>
<tr>
<td>Tropicana Prm 64 oz.</td>
<td>-9.413</td>
<td>7.152</td>
</tr>
<tr>
<td>Tropicana Prm 96 oz.</td>
<td>-19.669</td>
<td>5.383</td>
</tr>
<tr>
<td>Florida 96 oz.</td>
<td>-2.527</td>
<td>1.896</td>
</tr>
<tr>
<td>-Log(price)</td>
<td>-2.336</td>
<td>1.675</td>
</tr>
<tr>
<td>Display</td>
<td>0.387</td>
<td>0.214</td>
</tr>
<tr>
<td>-Log(price)-ethnic</td>
<td>4.165</td>
<td>0.013</td>
</tr>
<tr>
<td>-Log(price)-hval150</td>
<td>-1.363</td>
<td>0.013</td>
</tr>
<tr>
<td>Ethnic</td>
<td>8.750</td>
<td>0.013</td>
</tr>
<tr>
<td>Hval150</td>
<td>-1.669</td>
<td>0.013</td>
</tr>
<tr>
<td>Dvrtn</td>
<td>-0.051</td>
<td>0.013</td>
</tr>
<tr>
<td>Age60</td>
<td>1.367</td>
<td>0.013</td>
</tr>
<tr>
<td>Hlarge</td>
<td>0.013</td>
<td>0.013</td>
</tr>
<tr>
<td>$\sigma_d$</td>
<td>0.103</td>
<td>0.013</td>
</tr>
</tbody>
</table>

GMM objective value 395.614

Notes. Parameters represent $V_{ist}$ 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.
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.

The term $\Delta$ is the estimated margin and $\hat{\Phi}$ is the covariance matrix of the observed margins. The minimum distance criterion corresponding to both models is 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

$$\epsilon_{ij}^p = p_{ij} \left( \int \frac{\partial P(C_{ijst} = 1, I_{ist} = 1)}{\partial p_{ijst}} \phi(\Lambda) \, d\Lambda \right)$$

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

$$\epsilon_{ij}^q = p_{ij} \left( \int \frac{\partial E(x_{ijst} | C_{ijst} = 1, I_{ist} = 1)}{\partial p_{ijst}} \phi(\Lambda) \, d\Lambda \right)$$

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

---

Table 3

<table>
<thead>
<tr>
<th>Product</th>
<th>Mean</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minute Maid 64 oz.</td>
<td>1.715</td>
<td>0.625</td>
</tr>
<tr>
<td>Minute Maid 96 oz.</td>
<td>0.932</td>
<td>0.323</td>
</tr>
<tr>
<td>Dominick’s 64 oz.</td>
<td>2.352</td>
<td>0.920</td>
</tr>
<tr>
<td>Tropicana Prm 64 oz.</td>
<td>1.450</td>
<td>0.555</td>
</tr>
<tr>
<td>Tropicana SB 64 oz.</td>
<td>1.651</td>
<td>0.670</td>
</tr>
<tr>
<td>Tropicana Prm 96 oz.</td>
<td>0.839</td>
<td>0.280</td>
</tr>
<tr>
<td>Florida Orange 64 oz.</td>
<td>1.751</td>
<td>0.633</td>
</tr>
</tbody>
</table>

---

$^9$ In predicting margins out-of-sample, it is important to control for the structural error $\xi_{ist}$. From our estimation, for each week, we have estimates of $\xi_{ist}$ for the 30 stores in our sample. When computing margins for a store-week out-of-sample, we integrate over the empirical distribution of $\xi$ 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

<table>
<thead>
<tr>
<th>Brand</th>
<th>Size (oz.)</th>
<th>Conditional share (%)</th>
<th>Average price</th>
<th>Average cost</th>
<th>True % margins</th>
<th>Predicted % margins¹</th>
<th>Correlations of true and pred: margins²</th>
<th>Predicted % margins</th>
<th>Correlations of true and pred: margins²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minute Maid</td>
<td>64</td>
<td>21.25</td>
<td>$2.25</td>
<td>$1.69</td>
<td>23.83</td>
<td>50.90</td>
<td>0.3688</td>
<td>41.99</td>
<td>0.5757</td>
</tr>
<tr>
<td>Minute Maid</td>
<td>96</td>
<td>4.35</td>
<td>$4.11</td>
<td>$2.98</td>
<td>26.42</td>
<td>38.43</td>
<td>0.4583</td>
<td>33.60</td>
<td>0.6214</td>
</tr>
<tr>
<td>Dominick’s</td>
<td>64</td>
<td>22.71</td>
<td>$1.65</td>
<td>$1.15</td>
<td>28.41</td>
<td>49.93</td>
<td>0.3794</td>
<td>40.56</td>
<td>0.6360</td>
</tr>
<tr>
<td>Tropicana Prm</td>
<td>96</td>
<td>21.87</td>
<td>$2.67</td>
<td>$1.88</td>
<td>28.41</td>
<td>53.84</td>
<td>0.3432</td>
<td>36.00</td>
<td>0.5737</td>
</tr>
<tr>
<td>Tropicana Prm</td>
<td>64</td>
<td>17.69</td>
<td>$2.38</td>
<td>$1.62</td>
<td>30.39</td>
<td>52.13</td>
<td>0.4223</td>
<td>34.46</td>
<td>0.6954</td>
</tr>
<tr>
<td>Tropicana Prm</td>
<td>96</td>
<td>8.12</td>
<td>$4.52</td>
<td>$3.26</td>
<td>27.20</td>
<td>51.70</td>
<td>0.2320</td>
<td>36.47</td>
<td>0.5986</td>
</tr>
<tr>
<td>Florida</td>
<td>96</td>
<td>3.97</td>
<td>$2.18</td>
<td>$1.48</td>
<td>31.80</td>
<td>31.94</td>
<td>0.4775</td>
<td>32.44</td>
<td>0.5021</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Brand</th>
<th>Size (oz.)</th>
<th>Conditional share (%)</th>
<th>Average price</th>
<th>Average cost</th>
<th>True % margins</th>
<th>Predicted % margins</th>
<th>Correlations of true and pred: margin</th>
<th>Predicted % margins</th>
<th>Correlations of true and pred: margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minute Maid</td>
<td>64</td>
<td>21.16</td>
<td>$2.23</td>
<td>$1.69</td>
<td>23.61</td>
<td>48.87</td>
<td>0.2403</td>
<td>49.53</td>
<td>0.3894</td>
</tr>
<tr>
<td>Minute Maid</td>
<td>96</td>
<td>4.02</td>
<td>$4.08</td>
<td>$2.99</td>
<td>26.03</td>
<td>36.98</td>
<td>0.2022</td>
<td>31.48</td>
<td>0.3621</td>
</tr>
<tr>
<td>Dominick’s</td>
<td>64</td>
<td>25.33</td>
<td>$1.64</td>
<td>$1.15</td>
<td>28.03</td>
<td>48.86</td>
<td>0.2885</td>
<td>45.33</td>
<td>0.5237</td>
</tr>
<tr>
<td>Tropicana Prm</td>
<td>64</td>
<td>19.75</td>
<td>$2.66</td>
<td>$1.88</td>
<td>27.95</td>
<td>51.73</td>
<td>0.2439</td>
<td>35.81</td>
<td>0.5685</td>
</tr>
<tr>
<td>Tropicana Prm</td>
<td>96</td>
<td>7.47</td>
<td>$4.50</td>
<td>$3.26</td>
<td>26.94</td>
<td>55.18</td>
<td>0.0658</td>
<td>34.42</td>
<td>0.3355</td>
</tr>
<tr>
<td>Florida</td>
<td>96</td>
<td>4.53</td>
<td>$2.17</td>
<td>$1.48</td>
<td>31.41</td>
<td>29.30</td>
<td>0.2893</td>
<td>29.16</td>
<td>0.3053</td>
</tr>
</tbody>
</table>

Minimum distance criterion 1.2975 0.3083

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

<table>
<thead>
<tr>
<th>Effect of change in price of:</th>
<th>On the demand for:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM64</td>
<td>MM96</td>
</tr>
<tr>
<td>MM96</td>
<td>0.0033</td>
</tr>
<tr>
<td>Dom64</td>
<td>0.3480*</td>
</tr>
<tr>
<td>TRPrm64</td>
<td>0.0091</td>
</tr>
<tr>
<td>TRSB64</td>
<td>0.1010*</td>
</tr>
<tr>
<td>TRPrm96</td>
<td>0.1450*</td>
</tr>
<tr>
<td>FLR</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>Effect of change in price of:</th>
<th>On the unconditional probability of choice of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM64</td>
<td>MM96</td>
</tr>
<tr>
<td>MM96</td>
<td>-1.6200</td>
</tr>
<tr>
<td>Dom64</td>
<td>0.0033</td>
</tr>
<tr>
<td>TRPrm64</td>
<td>0.3480</td>
</tr>
<tr>
<td>TRSB64</td>
<td>0.0091</td>
</tr>
<tr>
<td>TRPrm96</td>
<td>0.1010</td>
</tr>
<tr>
<td>FLR</td>
<td>0.1450</td>
</tr>
</tbody>
</table>

No purchase probability

0.0003 0.0034 0.0005 0.0017 0.0011 0.0006 -2.1900 0.0040
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 test (Morrison 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; and \( \hat{W}_B \) 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 \( \mu_Q \) and \( \mu_B \) 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

\[ PD_{\mu} = \int \left( 1 + \alpha \int \left( 1 + \sum_{k}^n \frac{\mu(C_{\mu} = 1 | I_{\mu} = 1)}{\ln[\mu(I_{\mu} = 0)]} + \frac{\mu(C_{\mu} = 1 | I_{\mu} = 0)}{\ln[\mu(I_{\mu} = 0)]} \right) \right) d\Omega \]

\[ SD_{\mu} = 1 - PD_{\mu} \]

---

11 The model structure implies that the expected quantity elasticity is: \( e_{Qij} = -1 + f \alpha \mu_i \ln[\Pr(I_{ij} = 0)] \) [where ]\( \alpha = \sum_{k}^n \frac{\Pr(C_{ij} = 1 | I_{ij} = 1)}{\ln[\Pr(I_{ij} = 0)]} \). Therefore, low unconditional purchase probabilities, \( \Pr(C_{ij} = 1 | I_{ij} = 1) \), imply quantity elasticities close to 1.

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:
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

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.

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

<table>
<thead>
<tr>
<th>Effect of change in price of:</th>
<th>On the market share of:</th>
<th>No purchase probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM64</td>
<td>-0.0405*</td>
<td>0.0057</td>
</tr>
<tr>
<td>MM96</td>
<td>0.0237</td>
<td>0.0373*</td>
</tr>
<tr>
<td>Dom64</td>
<td>0.2120*</td>
<td>0.0110</td>
</tr>
<tr>
<td>TRPm64</td>
<td>0.2580*</td>
<td>0.0224</td>
</tr>
<tr>
<td>TRS864</td>
<td>0.3330*</td>
<td>0.0154</td>
</tr>
<tr>
<td>TRPrm96</td>
<td>0.1190*</td>
<td>0.0003</td>
</tr>
<tr>
<td>FLR</td>
<td>0.0011</td>
<td>0.0092</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>Brand</th>
<th>Size (oz.)</th>
<th>Elasticity based decomposition¹</th>
<th>Unit Sales based decomposition²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minute Maid 64</td>
<td>39.53</td>
<td>61.47</td>
<td>84.34</td>
</tr>
<tr>
<td>Minute Maid 96</td>
<td>32.94</td>
<td>67.06</td>
<td>96.70</td>
</tr>
<tr>
<td>Dominick’s 64</td>
<td>36.86</td>
<td>63.14</td>
<td>91.50</td>
</tr>
<tr>
<td>Tropicana Pm 64</td>
<td>37.22</td>
<td>62.78</td>
<td>97.29</td>
</tr>
<tr>
<td>Tropicana SB 64</td>
<td>32.66</td>
<td>67.34</td>
<td>95.56</td>
</tr>
<tr>
<td>Florida 96</td>
<td>31.52</td>
<td>68.48</td>
<td>98.68</td>
</tr>
<tr>
<td>Average for category</td>
<td>34.73</td>
<td>65.27</td>
<td>92.31</td>
</tr>
</tbody>
</table>

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).
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 \( \delta \) is a contraction mapping. That is, iteratively solving (12) converges to a vector \( \delta \) that uniquely reconciles \( \tilde{Q}_{ist} \) in Equation (11) with the average quantity per customer in the data, \( q_{ist} \).

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

\[
\tilde{Q}_{ist}(\delta; \Theta) = \int -\frac{y_{ist}}{\alpha_{ist}} \Pr(D_{ist} = 1 | I_{ist} = 1) \ln[P(I_{ist} = 0)]\phi(\Lambda) d\Lambda
\]

\[
= \int \frac{y_{ist}}{\alpha_{ist}} \sum_{k=1}^{j} e^{\delta_{j} + \Theta_{j} + \Omega_{j} \delta} \ln \left[ 1 + \sum_{k=1}^{j} e^{\delta_{j} + \Theta_{j} + \Omega_{j} \delta} \right] dF(\Omega).
\]

The iterative function \( g(\cdot) : \mathbb{R}^l \to \mathbb{R}^l \) (Equation (12)) is defined as:

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

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 \( \delta \);
(b) \( \frac{\partial g(\delta)}{\partial \delta} \geq 0 \quad \forall \, r, j \); and,
(c) \( \sum_{r=1}^{l} \frac{\partial g(\delta)}{\partial \delta_r} < 1 \).

The function is continuous by construction. To prove (b), we first show that \( \frac{\partial [g(\delta)]}{\partial \delta_r} \geq 0 \), and then that \( \frac{\partial [g(\delta)]}{\partial \delta_r} \geq 0 \), \( \forall \, r \neq j \).

(b.1) To show that \( \frac{\partial g(\delta)}{\partial \delta_j} \geq 0 \), note that

\[
\frac{\partial g(\delta)}{\partial \delta_j} = \Pr(C_i = j | I_i = 1) * [1 - \Pr(C_i = j | I_i = 1)],
\]

and that

\[
\frac{\partial}{\partial \delta_j} \left( \int 1 + \sum_{k=1}^{j} e^{\delta_j + \Theta_{j} + \Omega_{j}} \right) = \Pr(C_i = j, I_i = 1).
\]

Also note that at a given guess of the parameter vector, \( \alpha \) is known and is fixed. Hence,

\[
\frac{\partial [g(\delta)]}{\partial \delta_j} = 1 - \frac{1}{\tilde{Q}(\cdot)} \int \frac{y_{ij}}{\alpha_{ij}} \Pr(C_i = j | I_i = 1) \Pr(C_i = j, I_i = 1) \cdot \ln[\Pr(I_i = 0)] \cdot [1 - \Pr(C_i = j | I_i = 1)] dF(\Omega).
\]

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

\[
[\Pr(C_i = j, I_i = 1) - \ln(\Pr(I_i = 0))] [1 - \Pr(C_i = j | I_i = 1)] \leq - \ln(\Pr(I_i = 0)),
\]

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 \( \frac{\partial g(\delta)}{\partial \delta_r} \geq 0 \), \( r \neq j \), note first that

\[
\frac{\partial}{\partial \delta_r} \left( \frac{1}{\sum_{k=1}^{j} e^{\delta_j + \Theta_{j} + \Omega_{j}}} \right) = \frac{1}{\tilde{Q}(\cdot)} \int \frac{y_{ir}}{\alpha_{ir}} \Pr(C_i = j | I_i = 1) \Pr(C_i = r | I_i = 1) \cdot \ln(\Pr(I_i = 0)) dF(\Omega).
\]

\[
\geq 0.
\]

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

\[
\sum_{r=1}^{l} \frac{\partial g(\delta)}{\partial \delta_r} = 1 - \frac{1}{\tilde{Q}(\cdot)} \int \frac{y_{ir}}{\alpha_{ir}} \Pr(C_i = j | I_i = 1) dF(\Omega),
\]

where,

\[
f_i = \Pr(C_i = j, I_i = 1) - \Pr(I_i = 0) \cdot [1 - \Pr(C_i = j | I_i = 1)] + \ln(\Pr(I_i = 0)) \cdot \Pr(C_i = r | I_i = 1).
\]
Comparing the numerator of the second term in (A4) to line 1 of Equation (A0), we can see that for \( \sum_{j=1}^{J} \frac{\partial g_i}{\partial \delta} < 1 \), it is equivalent to prove that: \( f_i < -\ln[\Pr(i = 1)] \). This reduces to showing that
\[
\Pr(i = 1) \ast [\Pr(i = 1) + \ln(1 - \Pr(i = 1))] < 0.
\]
This holds since \( \Pr(i = 1) \in [0, 1] \).
Hence, \( g() \) is a contraction mapping, and therefore, iteratively solving (A1) will converge to a unique vector \( \delta \).

B. Method of Simulated Moments Procedure
Recall from §3 that we construct moment conditions based on the mean-independence assumption \( E[\xi_{it}Z_{it} | Z_{it}] = 0 \), where \( Z_{it} \) 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, \( \Theta^{GMM} \), that set \( M(\Theta) \) as close as possible to zero. We estimate all parameters by minimizing the GMM objective function, \( \Omega(\Theta) = M(\Theta)WM(\Theta) \). 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 (noneyconomic) 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 consider the case that the unobserved brand characteristics, \( \xi_{it} \), 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 Dominic 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 \( \text{Uniform}(5, 20) \). Average quantities per consumer for the 90 weeks were generated by integrating over the expected demands of consumers who are allowed to make multi-unit purchases. We report the results in Table A1 and 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. 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 multi-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, \( \alpha_r \), 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 \( \text{Uniform}(2, 5) \); 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

...
% change in category purchase probability:

<table>
<thead>
<tr>
<th></th>
<th>True</th>
<th>Model</th>
<th>True</th>
<th>Model</th>
<th>True</th>
<th>Model</th>
<th>True</th>
<th>Model</th>
<th>True</th>
<th>Model</th>
<th>True</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand 1 intercept</td>
<td>-2.000</td>
<td>-3.646</td>
<td>-2.000</td>
<td>-3.227</td>
<td>-2.000</td>
<td>-2.926</td>
<td>-2.000</td>
<td>-2.699</td>
<td>-2.000</td>
<td>-2.515</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 2 intercept</td>
<td>-4.000</td>
<td>-5.650</td>
<td>-4.000</td>
<td>-5.230</td>
<td>-4.000</td>
<td>-4.928</td>
<td>-4.000</td>
<td>-4.700</td>
<td>-4.000</td>
<td>-4.516</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% change in conditional choice probability:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% change in category purchase probability:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Data for this table: 190 weeks of data generated from logit demand model with Brand 1 prices \( \sim \text{iid } N(0, 0.1) \). "True" = Logit; "Model" = Discrete/continuous.

(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
where tity to this model is: parameter tional quantity: In contrast, our proposed specification has expected condi-

tions are (Equation 11, p. 33 in AAG): AAG framework.

tion 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):

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

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

\[ E[Q_{\text{discrete}}] = \sum_{q=5}^{\infty} [q * p_0(q)]. \]

In contrast, our proposed specification has expected conditional quantity:

\[ E[Q_{\text{continuous}}] = \frac{\gamma}{p_j} \left[ \alpha/\gamma - \ln(p_i) + \mu (\Gamma + \ln(1/\Pr(j))) \right], \]

where \( \Gamma \) 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 \( \alpha/\gamma \) and \( \gamma \). For all computations, \( \mu = 1 \), and “\( \infty \)” = 50 (in the summation on \( E[Q_{\text{discrete}}] \)). The results are given in Table A5. We see that 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


---

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

<table>
<thead>
<tr>
<th>Due to a 1% increase in price of:</th>
<th>True</th>
<th>Model</th>
<th>True</th>
<th>Model</th>
<th>True</th>
<th>Model</th>
<th>True</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand 1 intercept</td>
<td>−1.00</td>
<td>2.418</td>
<td>−1.00</td>
<td>1.707</td>
<td>−1.00</td>
<td>1.298</td>
<td>−1.00</td>
<td>1.098</td>
</tr>
<tr>
<td>Brand 2 intercept</td>
<td>−1.00</td>
<td>2.421</td>
<td>−1.00</td>
<td>1.689</td>
<td>−1.00</td>
<td>1.305</td>
<td>−1.00</td>
<td>1.069</td>
</tr>
<tr>
<td>Log(Price)</td>
<td>−2.00</td>
<td>3.610</td>
<td>−2.500</td>
<td>3.778</td>
<td>−3.00</td>
<td>4.152</td>
<td>−3.500</td>
<td>4.586</td>
</tr>
<tr>
<td>Mean conditional quantity</td>
<td>4.047</td>
<td>3.102</td>
<td>2.590</td>
<td>2.216</td>
<td>1.919</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Notes.

190 weeks of data generated from logit demand model with Brand 1 prices = uniform(2, 5), Brand 2 prices = uniform(2, 5), \( y = 25 \), \( \xi = \text{iid N}(0, 0.1) \). "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 overstat-

ing 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):

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

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

\[ E[Q_{\text{discrete}}] = \sum_{q=5}^{\infty} [q * p_0(q)]. \]

In contrast, our proposed specification has expected conditional quantity:

\[ E[Q_{\text{continuous}}] = \frac{\gamma}{p_j} \left[ \alpha/\gamma - \ln(p_i) + \mu (\Gamma + \ln(1/\Pr(j))) \right], \]

where \( \Gamma \) 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 \( \alpha/\gamma \) and \( \gamma \). For all computations, \( \mu = 1 \), and “\( \infty \)” = 50 (in the summation on \( E[Q_{\text{discrete}}] \)). The results are given in Table A5. We see that 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.

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

<table>
<thead>
<tr>
<th>( \alpha/\gamma = 3.0 )</th>
<th>( \gamma = 1.0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma )</td>
<td>( E[Q_{\text{continuous}}] )</td>
</tr>
<tr>
<td>1.0</td>
<td>0.9677</td>
</tr>
<tr>
<td>1.1</td>
<td>1.0644</td>
</tr>
<tr>
<td>1.2</td>
<td>1.1612</td>
</tr>
<tr>
<td>1.3</td>
<td>1.2580</td>
</tr>
<tr>
<td>1.4</td>
<td>1.3547</td>
</tr>
<tr>
<td>1.5</td>
<td>1.4515</td>
</tr>
<tr>
<td>1.6</td>
<td>1.5483</td>
</tr>
<tr>
<td>1.7</td>
<td>1.6450</td>
</tr>
<tr>
<td>1.8</td>
<td>1.7418</td>
</tr>
<tr>
<td>1.9</td>
<td>1.8386</td>
</tr>
<tr>
<td>2.0</td>
<td>1.9353</td>
</tr>
</tbody>
</table>


