TECHNICAL APPENDIX

"Accounting for Primary and Secondary Demand Effects with Aggregate Data"

Harikesh Nair, Jean-Pierre Dubé, Pradeep Chintagunta

Graduate School of Business
University of Chicago

Abstract

We report results from additional simulation studies and robustness checks related to the proposed model in this technical appendix. Study 1 presents a simulation study that shows that the proposed model performs reasonably well when applied to data generated from a Poisson model. Study 2 assesses the importance of stock-piling considerations for our data using household data for the refrigerated juice category. Study 3 discusses our exogeneity assumption of market-sizes, and study 4 presents sensitivity analysis related to using display as an instrument for shelf prices.

1. An alternative form of consumer behavior with brand and quantity choices

To demonstrate robustness of the proposed specification, we generate data from a different (non-economic) model of consumer choice behavior. We generate the data from the brand-choice-then-quantity model that is analogous to Dillon and Gupta (1996). In this model, purchase incidence is modeled as a logit and the quantity, conditional on purchase, is modeled as a truncated Poisson. For simplicity, we consider the case in which a homogenous set of consumers choose between 1 product and an outside good. In this model, the probability that the product is purchased, d = 1, is:

$$\Pr(d=1)=\frac{e^{\alpha-\beta p}}{1+e^{\alpha-\beta p}}$$

Conditional on purchase, the quantity chosen is distributed a truncated Poisson:

$$\Pr(Q=q|d=1,q>0)=\frac{e^{-\lambda}\lambda^{q}}{(1-e^{-\lambda})q!}$$

The expected conditional quantity is therefore,

$$E(Q|d=1,Q>0)=\frac{e^{-\lambda}}{(1-e^{-\lambda})}$$

Then, the expected per consumer demand is:

$$E[Q]=\Pr(d=1)E(Q|d=1,Q>0)=\frac{e^{\alpha-\beta p}}{1+e^{\alpha-\beta p}}\frac{e^{-\lambda}}{(1-e^{-\lambda})}$$

Let y denote expenditures. To capture the dependence of expected quantities on prices (p) and expenditures (y), we pick the specification $\lambda = \exp(\ln(y) - p)$. We first generate N = 200 expenditures as uniform(5,20) and prices as uniform(2,5). We then generate expected per-consumer demands from the above brand-choice-then-quantity model for various values of the parameters, and then run the proposed model on that data. Table 1 shows the results averaged across 100 replications.

As we see from the results in table 1, the proposed model does a reasonable job of recovering the choice probabilities and conditional quantities from the data. We take the performance of the model in the above situation as an indication that the model does not strongly a priori restrict choice probabilities and conditional quantities.

2. Stock-piling and inventories

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, Imai and Keane 2003 and Hendel and Nevo 2003), incorporating such sophisticated choice behavior at the aggregate level is beyond the scope of the current analysis.

To assess the importance of stock-piling considerations for our data, we looked at household data for this category. We obtained data for households making shopping trips in Denver between January 1993 and May 1995. While these data do not match the Chicago market or the precise time period, we have no reason a priori to expect Dominick's consumers to purchase these categories in different proportions. We report results from running brand-choice models with and without including an inventory variable as a demand-shifter. We measure inventory as the total volume in ounces of orange juice (to distinguish pack sizes) depleted over time using the average consumption rate for the household. The results are presented in Table 2.1. In terms of model parameters, we see that the inventory coefficient is significant (but of the wrong sign) and its inclusion leads to slightly different parameter values on other coefficients.

To assess the economic significance of inventories, we compute several own-price elasticities using the results from the inventory model. We compute elasticities at the observed inventory levels, setting inventories to zero for all trips, and setting inventory to each household's maximum observed level (i.e. if a household's highest inventory over time is 100 ounces, then we set inventory to 100 ounces for all storetrips for that household). We find that elasticities are virtually unchanged across specifications (see Table 2.2). The results also hold after accounting for heterogeneity (i.e. random coefficients). Hence, this does not appear to matter in terms of measuring price elasticity *for our category*. Of course, we do not generalize the minor impact of inventories that we find here to other categories.

3. Store choice and endogenous market sizes

For the current analysis, we assume that the market size (i.e. the number of people visiting the store each week 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.

We explore whether market-size (i.e. number of trips to the store each week) is endogenous in the data. Specifically, we check whether store-traffic is affected by category level prices for the refrigerated juice category. For this purpose, for each DFF store, we consider the total number of households in its trading area. Assuming each household shops once per week, we can evaluate the share of the shopping-trip market that comes to the store each week (as store trips/households). We regress the share of the market so computed, on category level price indices, controlling for store-level characteristics and holiday effects. In essence, this a "market-share" model of store choice for the store in question, with the covariates being the weekly category price indices for all the categories in the DFF database. The results are presented in Table 3. We find that refrigerated juice prices are insignificant in driving store-traffic. This result also holds if we instead run a regression of log-traffic on these indices. Hence, we treat store-choice and market-size (N) as exogenous for these data.

4. Display as an instrument

We check the sensitivity of our results to using display as an instrument. Table 4 reports results from estimating the homogenous discrete/continuous model with and without including display as an instrument. We effectively find that the results from the model are relatively unchanged if we drop display as an instrument. We find that the parameters change very little, and that further, there is only a small loss in fit from dropping display. The underlying reason is that most of the variation in the endogenous variable, viz. prices, are explained by wholesale costs, and the incremental gain in explaining the price variation from adding display is much less. Finally, previous research has routinely rejected the hypothesis that displays correlate with ξ and, hence, are invalid instruments (e.g. Sudhir 2001 and Chintagunta et al. 2003). These studies use Hausman statistics to test the hypothesis that displays are exogenous. These studies fail to reject this hypothesis.

Table 1: *Simulation results from running the proposed model on data from brand-choice-Poisson model*

	True	Model	True	Model	True	Model	True	Model
$\beta = 3.5$	α = -3		α = -4		α = -5		α = -6	
$\Pr(d=1)*1e3$	0.0011	0.0012	0.3435	0.3814	0.1477	0.1802	0.4992	0.5421
$E(Q \mid d=1, Q>0)$	1.2966	1.1702	1.2867	1.1998	1.3116	1.1449	1.3084	1.2542
α = -5	$\beta = 2.5$		$\beta = 3$		$\beta = 3.5$		$\beta = 4$	
$\Pr(d=1)*1e3$	0.3932	0.3461	0.2325	0.2211	0.1637	0.1949	0.0823	0.1122
$E(Q \mid d=1, Q>0)$	1.2880	1.5976	1.3169	1.4720	1.3481	1.1976	1.2894	1.0103

Table 2.1: *Parameter results with and without including inventory*

	MODEL 1		MODEL 2	
	Parameter	*Std. error*	*Parameter*	*Std. error*
MM 64	-0.991	0.212	-1.322	0.170
MM 96	-1.024	0.211	-1.332	0.176
TR 64	0.314	0.245	-0.038	0.201
TR 96	-1.240	0.274	-1.587	0.224
Price	-85.047	5.839	-82.832	4.826
Feature	0.085	0.070	0.109	0.070
Display	1.977	0.380	2.104	0.403
Inventory			1.300E-03	1.000E-04

Table 2.2: *Effects of including inventory on price elasticities*

	OWN-PRICE ELASTICITY		
Brand	Observed inventory	Zero inventory	Max: inventory
MM 64	-2.88	-2.89	-2.85
MM 96	-2.88	-2.89	-2.86
TR 64	-3.32	-3.34	-3.26
TR 96	-3.72	-3.72	-3.71

Table 3: *Regression of share of traffic on store characteristics and category price indices*

		Share of market	
		Parameter	Std. Error.
Store characteristics	Holiday	-0.38	0.09
	Average income	-1.69	0.36
	Mean residential value	0.02	0.00
	Proportion of population of age > 60	-1.11	0.68
	Proportion of ethnic population	0.23	0.46
	Shopping index	2.09	0.37
	Distance to nearest Jewel	0.31	0.03
	EDLP	-0.04	0.01
Log Price Index	Analgesics	1.95	0.33
	Bath soap	0.58	0.13
	Beer	-5.12	0.79
	Bottled juices	-3.57	0.86
	Canned cooking soups	2.26	0.55
	Cereals	0.62	0.42
	Cigarettes	-1.66	0.80
	Cookies	-1.92	0.54
	Crackers	-0.13	0.77
	Dish detergent (liquid)	0.48	0.72
	Canned eating soups	-1.87	0.50
	Front-end candies	0.50	0.52
	Frozen dinners	1.51	0.39
	Frozen entrees	1.73	0.34
	Frozen juices	-0.74	0.34
	Fabric softeners	-1.24	0.96
	Grooming products	-1.15	0.42
	Laundry detergents	-1.10	0.82
	Non-sliced cheeses (shredded, party, etc.)	1.50	0.49
	Dish detergent (powder)	-0.92	0.50
	Canned salmon, crabs, etc	1.06	0.36
	Oatmeal	-2.83	0.66
	Paper towels	-3.11	1.24
	Refrigerated juices	**-0.20**	**0.16**
	Sliced cheeses	3.09	0.66
	Soft drinks	2.05	0.64
	Shampoos	0.70	0.79
	Snack crackers	1.65	0.59
	Soaps	1.38	1.39
	Toothbrushes	-4.60	1.09
	Canned tuna	-0.43	0.26
	Toothpastes	-0.83	0.50
	Bathroom tissues	-3.89	0.85
Constant		44.14	5.15
Number of observations		1486	
R^2		0.37	

Table 4: *Parameter estimates (homogenous case) with and without display as instruments*

Variable	With Display as instrument		Without Display as an instrument	
	Parameter	*t*-stat	Parameter	*t*-stat
MinuteMaid 64 oz	-3.147	-30.863	-3.124	-29.087
MinuteMaid 96 oz	-3.892	-35.535	-3.915	-33.772
Dominicks 64 oz	-3.874	-42.220	-3.859	-41.116
Tropicana Prm 64 oz	-2.464	-22.499	-2.480	-21.411
Tropicana SB 64 oz	-3.432	-32.955	-3.401	-31.218
Tropicana Prm 96 oz	-2.973	-25.989	-3.010	-24.709
Florida 96 oz	-6.075	-69.179	-6.040	-67.214
-Log(price)	2.614	28.767	2.531	25.646
Display	0.467	17.945	0.486	17.456
-Log(price)*ethnic	3.963	11.394	3.957	11.103
-Log(price)*hval150	-2.737	-25.379	-2.705	-24.915
Ethnic	5.116	15.362	5.023	14.745
Hval150	-1.809	-17.721	-1.815	-17.676
Drvtm	-0.039	-6.072	-0.040	-6.296
Age60	0.422	3.499	0.500	4.124
Hhlarge	0.064	2.140	0.074	-2.460
Objective function	1026.67		1053.69	