



**UCD GEARY INSTITUTE  
DISCUSSION PAPER  
SERIES**

**Coverage of Retail Stores and  
Discrete Choice Models of Demand:  
Estimating Price Elasticities and  
Welfare Effects**

January 2010

# Coverage of Retail Stores and Discrete Choice Models of Demand: Estimating Price Elasticities and Welfare Effects\*

Franco Mariuzzo<sup>†</sup>  
Geary Institute,  
University College Dublin,  
Ireland.

Patrick Paul Walsh<sup>‡</sup>  
SPIRe and Geary Institute,  
University College Dublin,  
Ireland.

Ciara Whelan  
School of Economics,  
University College Dublin, Ireland.

July 21, 2009

## Abstract

Consumers' choice set of products within stores can be limited. Akerberg and Rysman (2005) address this problem by modeling unobserved consumer preferences over products *and* retail stores, leading to augmented demand specifications. Having Carbonated Soft Drink product level data, where we observe products' store coverage, we are able to estimate their logit, nested logit and random coefficient logit specifications of demand in a structural model of equilibrium. Allowing for store coverage turns out to have a very significant impact on the estimated structural parameters and on the predictive power of the model. Taking these estimated structural parameters we perform a counterfactual whereby stores carry all products in the market. We find systematic increases in price elasticities and welfare in our new equilibrium. Competition in markets is more curtailed than normally assumed in structural models of industries.

KEYWORDS: Carbonated Soft Drinks, differentiated products, discrete choice, store coverage, structural model, price elasticities, welfare.

*JEL classifiers: L11 and L62.*

---

\*This paper was presented at the Harvard University IO group, the Industrial Organization Society meeting in Chicago and the Industry Group in the L.S.E.. We thank seminar participants for comments. In addition we thank John Asker, Agar Brugiavini, Peter Davis, Gautam Gowrisankaran, Katherine Ho, Julie Mortimer, Ariel Pakes, Mark Schankerman, Laura Spierdijk and John Sutton for detailed comments on earlier drafts. Finally we thank Daniel Akerberg and anonymous referees for extremely detailed comments on the paper.

<sup>†</sup>Mariuzzo gratefully acknowledges financial support from Marie Curie Excellence Grant MEXT-CT-2006-042471.

<sup>‡</sup>Corresponding authors: franco.mariuzzo@ucd.ie; ppwalsh@ucd.ie.

# 1 Introduction

The ability to estimate demand systems for differentiated products has become a core part of New Empirical Industrial Organization.<sup>1</sup> Theoretical innovations allow us to estimate demand for differentiated products at the product rather than at the consumer level. This is done by embedding a certain structure into consumer utility, such as the presence of unobserved heterogeneity in consumer taste for products around mean utilities using standard logit errors. Utilizing data on prices, market shares and product characteristics (observable and unobservable to the econometrician), logit and nested logit discrete-choice models can be evaluated using the methodology of Berry (1994). In addition, if data on individual characteristics (observable and unobservable to the econometrician) is accounted for, random coefficients, as in Berry, Levinsohn, and Pakes (1995) and Nevo (2001), can be estimated.

However, as noted by Akerberg and Rysman (2005), one should worry about the properties of the commonly used standard logit errors, since these can adversely affect substitution patterns and welfare outcomes. Logit errors assume the dimension of unobservable characteristic space to expand proportionally to the number of products. This may not be realistic for a consumer in a store, since stores do not tend to carry all available products in the market. This is confirmed by Pakes (2003), who shows that computer retailers typically display only a subset of the total number of computer models available. Thus, one important source of unobserved consumer taste heterogeneity is induced by product availability inside stores.<sup>2</sup> Yet, most empirical estimations of demand systems for differentiated products do not allow heterogeneous availability of products to enter consumer utility.

Another property of standard logit errors is that all products are equidistant from each other in unobserved characteristic space, when in reality consumer distance to products, due to store coverage, is expected to be idiosyncratic.<sup>3</sup> We address both of these issues in this paper.

Following Akerberg and Rysman (2005) we adjust logit errors to reflect

---

<sup>1</sup>See Akerberg et al. (2007), and Perloff et al. (2007) for a good literature review.

<sup>2</sup>Akerberg and Rysman (2005) list many examples where one can see cross-sectional variation in the number of available products in the data [Berry and Waldfogel (1999), Crawford (2000), Arcidiacono (2005), and Rysman (2004)], or time series variation [Berry, Levinsohn, and Pakes (1995), Bresnahan, Stern, and Trajtenberg (1997), Petrin (2002), and Crawford and Shum (2005)]. Nevo (2001), and Town and Liu (2003) deal with both.

<sup>3</sup>This would not be true in familiar theoretical location games such as the linear city of Hotelling (1929) and D'Aspremont et al. (1979), or the circular city of Salop (1979). Heterogeneity in consumer tastes (transportation costs) can be modelled empirically, see Davis (2006).

possible product congestion in stores, i.e. products competing for limited available shelf space, by allowing different indirect utility specifications to represent idiosyncratic unobserved consumer preferences over both retail stores and products.<sup>4</sup> We use scanner data on 157 products in retail Carbonated Soft Drinks (CSD) over 28 bimonthly periods where retail store coverage *is* observed at the product level, to estimate their derived logit, nested logit and random coefficients logit specifications in a structural model of equilibrium.

As Akerberg and Rysman’s (2005) Monte Carlo study shows, not allowing for congestion in stores can lead to biases in estimates of price elasticities, as well as inaccurate welfare effects. Our analysis based on observable retail product coverage, will provide us with the true structural parameters and will lead to more realistic outcomes in terms of price elasticities and welfare than assuming standard logit errors. Allowing for store coverage at the product level is shown to have the predicted significant impact on our jointly estimated demand and cost systems.

To understand the bias created by standard logit errors, and to further link our results to Akerberg and Rysman (2005), we impose full-coverage (no congestion) in our structural model of equilibrium and evaluate its effect on price elasticities and welfare. This counterfactual exercise highlights the potential distortions that could arise from assuming standard logit errors when congestion inside stores is actually a feature of consumer choice. All else equal, markets are found to be more competitive by design when one uses standard logit errors as the building block to allow for aggregation over unobserved heterogeneity in consumer taste at the product level.

The remainder of the paper is organized as follows. Section 2 summarizes our data and discusses the Akerberg and Rysman (2005) proxies for store coverage. Section 3 describes our three demand specifications and a supply side model. It also outlines our counterfactual and derives consumer welfare for the three specifications. Section 4 outlines our estimations and presents our empirical results. We make our conclusions in section 5.

---

<sup>4</sup> Another approach, advocated by Berry and Pakes (2007), Bajari and Benkard (2005), and Song (2007), would be to eliminate idiosyncratic logit errors from the model. The Akerberg and Rysman (2005) approach is to keep logit errors in the model but allow them to be more flexible. The flexible logit error imposes no additional computational burden but does require more product level data such as store coverage. As we will see, using a proxy based on the number of products may not be of much benefit.

## 2 Data

AC Nielsen collated a panel database of all CSD products distributed throughout 12,000 Irish retail stores. The database provides 28 bimonthly periods spanning October 1992 to March 1997 for 157 products (brands), identified for 13 firms and 40 product characteristics within the particular business of retail CSD. The data record the retail activities of both Irish and foreign owned products/firms selling throughout stores in the Irish retail sector.

The retail market for Carbonated Soft Drinks in Ireland is broadly similar in structure to that of the US. In 1997, the top two firms collectively account for 73 per cent of the Irish market and 75 per cent of the US retail market. Inequality in retail sales, as measured by the Gini coefficient, was 0.72 in Ireland and 0.68 in the US. Like the US the Irish market is heavily branded. There are differences between Ireland and the US that are typical of European CSD markets. These differences are highlighted in case studies of several countries in Sutton (1991).

We have product level information on:

- Product characteristics such as, flavors (Cola, Orange, Lemonade and Mixed Fruit), packaging types (Cans, Standard Bottle, 1.5 Liter, 2 Liter and Multi-Pack of Cans) and sweeteners (Diet and Regular).<sup>5</sup>

Allowing for flavor segments (Cola, Orange, Lemonade and Mixed Fruit) is standard in the analysis of CSD, (Sutton, 1991). To understand why packaging format is recognized as a crucial feature of this market, we graph the seasonal cycles of CSD sales by packaging type in Figure 1. The industry has introduced different packaging to satisfy different consumer needs (Walsh and Whelan, 1999). For example, cans peak in the summer months of June and July, when the weather is warmer. In contrast, 2 Liter bottle sales peak over the winter months of December and January, the festive season. Packaging by time dummies will be an important way to control for the impact of weather variations on the sales of segments defined by packaging. In the cost function, packaging by time dummies will control for plastic (2-Liter), glass (standard) and aluminium (cans) input costs that can change over-time.

- Price in Irish pounds per liter (weighted average of individual product

---

<sup>5</sup>The number of product characteristics was very stable throughout the period of this study. We take the emergence of such segments as an outcome of history. If in CSD consumers were fully mobile across segments and advertising was very effective, the market would evolve to be dominated by one segment. Taste structures and advertising outlays, amongst other factors, have driven the current day segmentation of the market by product attribute, see Sutton (1998).

prices across all stores selling the product, weighted by product sales share within stores).<sup>6</sup>

- Quantity sold (in thousands of liters).
- Sales value (in thousands of Irish pounds).

Details on the number of products, number of firms, prices, unit sales, and revenue shares by product segments are outlined in Table 1. We note that unit sales and revenue shares can be very different when one compares Cans to 2-Liters, for example.

- Inventory stocks (forward and back room) represent time to stock-out (expressed in months) of product sales for the bimonthly period. Table 2 details inventory levels by segments and time, averaged over products. A demand side interpretation would be that inventories reflect unanticipated slumps in demand. A supply side view is that inventories are a way of managing delivery costs.<sup>7</sup> We see clearly that segments depending more on large numbers of impulse buys/small stores (indicated by the cans and standard packaging format), use inventories more. Here, delivery costs are clearly higher and can be reduced with the use of inventories. The use of inventories is pretty stable over time and stock outs are not a feature of the data.
- Store coverage based on counts of stores that the product is in, weighted by store size in terms of overall CSD turnover. Table 2 also documents store coverage averaged over products within segments over time, and clearly shows that it varies over both time and segment.<sup>8</sup>

Akerberg and Rysman (2005), *AR* henceforth, do model unobserved consumer preferences over products by retail stores. Due to data unavailability, they resort to proxies of coverage. In creating the proxies the authors assume

---

<sup>6</sup>The average exchange rate over the period 1992-1997 is approximately 1.6 US dollar to one Irish pound.

<sup>7</sup>We refer to Deaton and Laroque (1992, 1996) for a theoretical aside on inventory models.

<sup>8</sup>The structure of the market has large companies placing products across most segments with varying coverage of stores inside segments depending on the product. Small companies can specialize in segments with a store type focus. The two large companies tend to face competition from different small independents within each segment. In each segment market size to sunk cost and the nature of price and non-price competition seem to limit the number of firms that can operate [see Mariuzzo, Walsh and Whelan (2003)]. The number of firms that operate in each segment is quite small. Yet, due to certain local taste characteristics, particularly in Orange and Mixed Fruit, small companies can fill a quality window and survive alongside large companies.

products equally cover a certain number of the total outlets so that, within a period, the store coverage for a representative product  $j$  can be written as,

$$R_j^{AR} = R_J^{AR} = \frac{R}{J}, \quad j = 1, 2, \dots, J,$$

where  $R$  indicates the total number of stores and  $J$  the total number of products marketed. They propose the following proxies for store coverage, written as a function of the total number of marketed products:

- i)  $R_J^{AR} = [\lambda_1/J + (1 - \lambda_1)]$
- ii)  $R_J^{AR} = J^{\lambda_2}$ ,

In specification (i) a value of  $\lambda_1$  approaching zero is indicative of what they call a “standard logit model”, i.e. the introduction of new products in the market is mirrored directly in new stores. This is as if the product dimension and the store dimension are able to track each other. On the other hand, a value of  $\lambda_1$  approaching one has an opposite effect that they name “pure congestion”. That is, new products have to compete for outlets with the existing products, thus providing a complete “crowding out” effect. In specification (ii) we have “pure congestion” if the parameter  $\lambda_2$  approaches minus one, and the “standard logit model” if  $\lambda_2$  goes to zero.

Contrary to  $AR$  we have information on store coverage at the product level in each time period. Coverage is measured as the proportion of stores each product is in, weighted by each store size in terms of CSD sales. This is denoted by  $r_{jt}$ , where  $0 < r_{jt} \leq 1$ . A value of one in  $r_{jt}$  indicates that the product is sold in all stores in period  $t$ . We now compare our actual measure of store coverage with the  $AR$  proxies.

We use our data to estimate  $AR$  store coverage proxies as in (i) and (ii). Our dependent variable,  $r_{jt}$ , has variability across products and time, whereas the explanatory variable (number of products) changes only over time,  $J_t$ . We suggest the following specification

$$r_{jt} = f(J_t; \lambda) \exp(\epsilon_{jt}), \quad (1)$$

where  $\epsilon_{jt}$  is the composite error, which we decompose into a time function  $h(t)$ , product random effects  $\kappa_j$ , and a pure idiosyncratic error  $u_{jt}$ ,

$$\epsilon_{jt} \equiv h(t) + \kappa_j + u_{jt},$$

with  $\kappa$  and  $u$  assumed to follow independent distributions.

Now, given that the main explanatory variables suggested in  $AR$  vary only over time, we average (1) with respect to the product dimension and take the natural log, to give the following relation,

$$\ln r_{.t} = \ln f(J_t; \lambda) + h(t) + \kappa_{.} + u_{.t}. \quad (2)$$

The function  $f(J_t; \lambda)$  may be, either  $[\lambda_1/J_t + (1 - \lambda_1)]$  under  $AR$  specification (i), or  $J_t^{\lambda_2}$  under  $AR$  specification (ii). The results of the nonlinear estimations of equation (2) under the functional forms (i) and (ii) are documented in Table 3. Both estimates of  $\lambda$  ( $\hat{\lambda}_1 = 1$  and  $\hat{\lambda}_2 = -0.9$ ) indicate product “congestion”. The predicted value of the  $AR$  proxy is depicted in the bottom panel of Figure 2. It delineates a down-trend, suggesting that product congestion has increased over time due to an increase in the number of products marketed (see the middle panel of Figure 2). Now we turn to the actual variability of our coverage variable (top part of Figure 2). The lower value of the boxes in Figure 2 is associated with the 25<sup>th</sup> percentile of the distribution, the upper value to the 75<sup>th</sup> percentile, and the line in the box captures the median value. Clearly the distribution of product coverage has increased in its variability over time, and this effect has been accompanied by a slight decrease in the median product coverage. We will use the estimated  $AR$  proxy and our product level, time varying, store coverage data in our empirical work. Given that the  $AR$  proxy varies only over time, we cannot expect it to capture any change in the within period product-level store coverage heterogeneity. This is something we want to remember when interpreting the role of the  $AR$  proxy in the estimations that will follow.

### 3 The Structural Model

In every period  $t$  a population of  $N_t$  individuals decides which of the available  $J_t + 1$  products to buy (where the +1 denotes an outside composite good). We stick to the existing discrete choice literature and assume individual preferences are represented in an indirect random utility function. In the next section we present alternative specifications of the demand side.

#### 3.1 Demand Specifications

##### 3.1.1 Logit Model

We define market size in terms of an average individual consumption of a can (220 ml) of soft carbonates per day. Consumer  $i$ ’s indirect utility,  $i = 1, 2, \dots, N_t$ ,



for product  $j$ ,  $j = 0, 1, \dots, J_t$ , in retail store  $r$ ,  $r = 1, 2, \dots, R_t$ , in period  $t$ ,  $t = 1, 2, \dots, T$ , can be expressed linearly as

$$u_{ijrt} = \alpha(y_{it} - p_{jt}) + \mathbf{x}_{jt}'\boldsymbol{\beta} + \xi_{jt} + \zeta_{ijt} + \sigma_r \varepsilon_{ijrt}, \quad (3)$$

where  $\mathbf{x}_{jt}$  is a column vector of  $k$ -observable product characteristics (inclusive of a constant),  $y_{it}$  is the income of individual  $i$  in period  $t$ ,  $p_{jt}$  denotes the price of the product in the period, and  $\xi_{jt}$  gathers any possible remaining unobservable (to the econometrician) product-period-level effects. As suggested by  $AR$ , the composite error term can be decomposed into a consumer  $i$ 's period-product-specific taste,  $\zeta_{ijt}$ , and consumer  $i$ 's period-product-retail store -specific taste,  $\varepsilon_{ijrt}$ . The parameter  $\sigma_r$  is a measure of the relative variance of the two random error variates.<sup>9</sup>

We define

$$\delta_{jt} \equiv \mathbf{x}_{jt}'\boldsymbol{\beta} - \alpha p_{jt} + \xi_{jt}$$

to be the mean utility over individuals and products.<sup>10</sup> As typical in the discrete choice literature, consumers can opt for an outside composite good whose utility can be represented by,

$$u_{i0rt} = \alpha y_{it} + \xi_{0t} + \zeta_{i0t} + \sigma_r \varepsilon_{i0rt}. \quad (4)$$

Now, assuming that  $\varepsilon_{ijrt}$  follows an *iid* type- $I$  extreme value distribution and that  $\zeta_{ijt}$  is distributed such that,  $\zeta_{ijt} + \sigma_r \varepsilon_{ijrt}$  is an *iid* type- $I$  extreme value (see Cardell, 1997), we can derive the following adjusted (for store coverage) logistic product level formula for market share as

$$s_{jt} = \frac{R_{jt}^{\sigma_r} \exp(\delta_{jt})}{\sum_{l=0}^{J_t} R_{lt}^{\sigma_r} \exp(\delta_{lt})}, \quad (5)$$

---

<sup>9</sup>The subscript  $r$  in the  $\sigma$  parameter has to be understood as a simple label, rather than a parameter that is store-specific (which here would indicate heteroskedasticity in the error component). The same interpretation will be on, for the  $\sigma_g$  and  $\sigma_p$  parameters that we will introduce below.

<sup>10</sup>It is useful to contrast the indirect utility in (3), with a more classical logit version

$$u_{ijrt} = \alpha y_{it} + \delta_{jt} + \varepsilon_{irt}.$$

The latter is basically a standard logit, with the only difference being that the idiosyncratic error,  $\varepsilon$ , is *iid* over stores rather than over products. Here, if one assumes stores to sell all products, the classical logit and standard logit would back out the same market share function, as confirmed in Figure 3. Needless to say, the classical logit is rather unrealistic, as it assumes stores to carry all products, while in reality they carry only an incomplete number. In addition to that, one can expect a certain degree of correlation for the same product sold across different stores. Hence, the indirect utility in (3) offers a more realistic representation. The underlying tree and derived market share function (demand function) are depicted in Figure 4.

where  $R_{jt}$  is the total number of retail stores carrying product  $j$  in period  $t$ , and the term  $\alpha y_{it}$  cancels out due to the way we have specified the utility.<sup>11</sup> The derivation of the term  $R_{jt}^{\sigma_r}$  is discussed in Appendix (6.A).

Next, if we divide the numerator and denominator of (5) by a function of the total number of stores in the economy in period  $t$ , i.e.,  $R_t^{\sigma_r}$ , we can re-express the market share as,

$$s_{jt} = \frac{r_{jt}^{\sigma_r} \exp(\delta_{jt})}{\sum_{l=0}^{J_t} r_{lt}^{\sigma_r} \exp(\delta_{lt})} \equiv \frac{\exp(\delta_{jt}^r)}{\sum_{l=0}^{J_t} \exp(\delta_{lt}^r)}, \quad (6)$$

where  $r_{jt} = R_{jt}/R_t$ , and the new definition of the mean utility  $\delta_{jt}^r$ , now accounts for the observable store coverage effect and relates to  $\delta_{jt}$  in the following way

$$\delta_{jt}^r \equiv \delta_{jt} + \sigma_r \ln r_{jt}.$$

Here, as is common, we set the mean utility of the outside composite good to zero,  $\delta_{0t} = 0$ . Moreover, we assume such a good to be carried in all stores in all periods, so that  $r_{0t} = r_0 = 1$ . If we take the natural log of each product's market share in (6) normalized with respect to the outside composite good we obtain the following reduced form equation which we estimate as our demand system for products,

$$\ln s_{jt} - \ln s_{0t} = \mathbf{x}_{jt}' \boldsymbol{\beta} + \sigma_r \ln r_{jt} - \alpha p_{jt} + \xi_{jt}.$$

So far we have been indirectly thinking of stores carrying only one product, while in reality stores carry portfolios of products, which have to compete for shelf space. The extension from single-product store to multi-product store is rather immediate. We leave to the reader to visualize how Figure (4) can be extended to this latter case. Once stores carry different products,  $r_{jt}$  becomes our effective measure of store coverage: the proportion of stores each product is in, weighted by stores market shares (measured in terms of sales). The next section extends this logit model to a nested logit framework.

### 3.1.2 Nested Logit Model

We follow *AR* and rewrite the indirect utility function as,

$$u_{ijrt} = \alpha y_{it} + \delta_{jt} + \zeta_{igt}^1 + \sigma_g (\zeta_{ijt}^2 + \sigma_r \varepsilon_{ijrt}), \quad (7)$$

and the corresponding function for the outside good as,

$$u_{i0rt} = \alpha y_{it} + \zeta_{i0t}^1 + \sigma_g (\zeta_{i0t}^2 + \sigma_r \varepsilon_{i0rt}),$$

---

<sup>11</sup>This specification justifies the fact that we do not have individual income data.

where, again, we normalize  $\delta_{0t}$  to zero and  $r_{0t} = r_0$  to one, and find that  $\alpha y_{it}$  cancels out when we subtract  $u_{i0rt}$  from  $u_{ijrt}$ .

Unlike equation (3), we now allow the random part of the utility to depend, each period  $t$ , on individuals taste for segment  $g$ ,  $g = 1, 2, \dots, G_t$ , individuals taste for products, and individuals product-store taste. Along with the random component we allow for a deterministic part which consists of an individual income, and the (product) mean utility  $\delta_{jt}$ , previously defined. Assuming  $\varepsilon_{ijrt}$  follows an *iid* type-I extreme value distribution,  $\zeta_{ijt}^2$  to have an *iid* distribution such that  $\zeta_{ijt}^2 + \sigma_r \varepsilon_{ijrt}$  is type-I extreme value, and  $\zeta_{igt}^1$  to have an *iid* distribution similar to  $\zeta_{igt}^1 + \sigma_g (\zeta_{ijt}^2 + \sigma_r \varepsilon_{ijrt})$ , we can back out nested logit market shares. Defining  $s_{j/g}$  to be the within group market share and  $s_g$  to be the segment market share, we derive the following expressions,

$$s_{j/g,t} = \frac{\exp\left(\frac{\delta_{jt}^r}{\sigma_g}\right)}{D_{gt}}, \quad s_{gt} = \frac{D_{gt}^{\sigma_g}}{\sum_g D_{gt}^{\sigma_g}}, \quad (8)$$

where  $\delta_{jt}^r$  is the mean utility defined in the previous section to account for the store coverage effect. The term  $D_{gt}$  captures the sum,

$$D_{gt} \equiv \sum_{l \in \mathcal{J}_{gt}} \exp\left(\frac{\delta_{lt}^r}{\sigma_g}\right), \quad (j, l \in \mathcal{J}_{gt}),$$

with  $\mathcal{J}_{gt}$  we denote the set of products marketed in segment  $g$  in period  $t$ .<sup>12</sup>

As is usual we express the nested logit specification as the product of two logits,

$$s_{jt} = s_{j/g,t} s_{gt}. \quad (9)$$

Using the Berry (1994) inversion, equation (9) can be rewritten as

$$\ln s_{jt} - \ln s_{0t} = \delta_{jt}^r + (1 - \sigma_g) \ln s_{j/g,t}, \quad (10)$$

In the next section we generalize the model further and allow retail store coverage to be embedded in a random coefficient nested logit model.

### 3.1.3 Random Coefficient Nested Logit Model

We now enrich the previous analysis by allowing random coefficients on the price (and income) parameters so that the indirect utility function can be cast as

$$u_{ijrt} = \alpha_i (y_{it} - p_{jt}) + \mathbf{x}_{jt}' \boldsymbol{\beta} + \xi_{jt} + \zeta_{igt}^1 + \sigma_g (\zeta_{ijt}^2 + \sigma_r \varepsilon_{ijrt}), \quad (11)$$

---

<sup>12</sup>This approach is a little restricted in the sense that one might prefer to model goods in one segment to crowd out (in terms of retail space) goods in the same segment more than goods in different segments.

with the utility for the outside good (with  $\delta_{0t} = 0$  and  $r_{0t} = r_0 = 1$ ) as,

$$u_{i0rt} = \alpha_i y_{it} + \zeta_{igt}^1 + \sigma_g (\zeta_{i0t}^2 + \sigma_r \varepsilon_{i0rt}),$$

where we have assumed the inside and outside market individual marginal utilities of income to be the same, so that in the process of aggregation the  $\alpha_i y_{it}$  term cancels out.

The random coefficient term can be written as

$$\alpha_i = \alpha + \sigma_p v_i^p,$$

where  $v_i^p$  is assumed to follow a standardized lognormal distribution.

As in Nevo (2000a) we write the products choice in terms of individual characteristics,

$$\mathcal{A}_{jt} [\boldsymbol{\delta}_t] = \{ [v_i^p, \zeta_{igt}^1 + \sigma_g (\zeta_{ijt}^2 + \sigma_r \varepsilon_{ijrt})] | u_{ijrt} \geq u_{ilrt}, \forall l = 0, 1, \dots, J_t; \quad t = 1, 2, \dots, T \}$$

where,

$$\boldsymbol{\delta}_t = [\delta_1, \delta_2, \dots, \delta_{J_t}].$$

Product market shares are obtained by integrating over the joint distributions of individual characteristics and stores,

$$s_{jt} \equiv E(s_{ijrt}) = \int_{\mathcal{A}_{jt}} dP [v^p, \zeta^1 + \sigma_g (\zeta^2 + \sigma_r \varepsilon)].$$

Next, conditioning the composite error  $(\zeta^1 + \sigma_g (\zeta^2 + \sigma_r \varepsilon))$  distribution on the joint distribution of  $v^p$ , we can simplify the above relation to,

$$s_{jt} \equiv E(s_{ijrt}) = \int_{\mathcal{A}_{jt}} dP [\zeta^1 + \sigma_g (\zeta^2 + \sigma_r \varepsilon) | v^p] dP(v^p).$$

Under the assumption that the composite error follows a type  $I$  extreme value distribution we can write individual probabilities of buying product  $j$  in period  $t$  as,

$$s_{ijt} = s_{i,j/g,t} s_{igt}, \quad (12)$$

where,

$$s_{i,j/g,t} = \frac{\exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} v_i^p}{\sigma_g}\right)}{D_{igt}}, \quad s_{igt} = \frac{D_{igt}^{\sigma_g}}{\sum_{s=0}^{G_t} D_{ist}^{\sigma_g}}, \quad (13)$$

with,

$$D_{igt} \equiv \sum_{l \in \mathcal{J}_{gt}} \exp\left(\frac{\delta_{jt}^r - \sigma_p p_{lt} v_i^p}{\sigma_g}\right), \quad (j, l \in \mathcal{J}_{gt}). \quad (14)$$

We use  $ns$  simulations to compute each product market share in period  $t$ :

$$s_{jt} = \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt}. \quad (15)$$

### 3.2 Supply side

We have detailed three alternative demand side models. For the cost (“supply”) side, we postulate constant product-period-specific marginal costs,  $c_{jt}$ , and deal with multi-product price setting firms. We assume the existence of a Nash Bertrand equilibrium, so that each period  $t$ , each firm  $f$ ,  $f = 1, \dots, F_t$ , maximizes the profit

$$\max_{p_{jt}} \sum_{l \in \mathcal{J}_{ft}} (p_{lt} - c_{lt}) s_{lt} (p_{1t}, p_{2t}, \dots, p_{J_t}) N_t, \quad (j, l \in \mathcal{J}_{ft}; f = 1, \dots, F_t),$$

where  $\mathcal{J}_{ft}$  is the set of products marketed by firm  $f$  in period  $t$ . Given the total number of firms available, we derive for each period a system of  $J_t$  first order conditions:

$$s_{jt} + \sum_{l \in \mathcal{J}_{ft}} (p_{lt} - c_{lt}) \frac{\partial s_{lt}}{\partial p_{jt}} = 0. \quad (16)$$

Next, if we define,

$$\Delta_t \equiv \begin{cases} -\frac{\partial s_{lt}}{\partial p_{jt}}, & \text{if products } (j, l \in \mathcal{J}_{ft}; f = 1, \dots, F_t) \\ 0, & \text{otherwise} \end{cases},$$

we can rewrite (16) in compact form as,

$$\Delta_t \equiv - \left[ \frac{\partial}{\partial \mathbf{p}_t} \mathbf{s}_t \right] \cdot \mathbf{O}_t, \quad (17)$$

where  $\mathbf{O}_t$  is period  $t$ 's firms' ownership matrix, and a dot indicates an element by element operation.<sup>13</sup>

The resulting pricing equations are

$$\mathbf{p}_t = \mathbf{c}_t + \underbrace{\Delta_t^{-1} \mathbf{s}_t}_{\mathbf{m}_t}. \quad (18)$$

which can be re-expressed at the product-period level as,

$$p_{jt} = c_{jt} + m_{jt}, \quad (19)$$

---

<sup>13</sup>The ownership matrix  $\mathbf{O}_t$ , is nothing more than a  $J_t \times J_t$  matrix of ones and zeros. The ones indicate products owned by the same firms.

where  $m_{jt}$  is the markup of product  $j$  in period  $t$ . Thus, the first order condition in (16) can be read as price equals marginal cost plus a mark-up, and given the primitives of the demand system one can use the pricing equation to back out marginal costs for each product.

As standard in the literature, we keep our model simple by postulating the existence of a linear relation between constant marginal costs and observable and unobservable (to the econometrician) product characteristics,

$$c_{jt} = \mathbf{w}_{jt}'\boldsymbol{\gamma} + \omega_{jt}. \quad (20)$$

Having defined the demand and cost functions, we now outline the consumer welfare in the models.

### 3.3 Consumer Welfare

Consumer welfare is defined as the monetary utility individuals receive by choosing one particular product. The marginal utility of income is constant and equal to:

$$\frac{\partial u}{\partial y} = \alpha,$$

in specifications [(3),( 7)], and is individual specific:

$$\frac{\partial u}{\partial y} = \alpha_i,$$

in specification (11).

Its inverse is used to convert consumers welfare in monetary terms. Small and Rosen (1981) have shown that in the case of *iid* extreme value idiosyncratic errors, and utilities linear in income, the expected consumer welfare has a closed form that, under our specifications, is written as follows:

1. **Logit:** Adjusting Train's (2003) formula to our notation we express consumer welfare as the monetary value of the maximum utility,

$$CW_{irt} = \frac{1}{\alpha} \max_{j \in J_t} u_{ijrt}.$$

However, due to the random composite error in the utility we are only able to formulate the expected consumer welfare,

$$E(CW_{irt}) = \frac{1}{\alpha} E \left[ \max_{j \in J_t} (\delta_{jt} + \zeta_{ijt} + \sigma_r \varepsilon_{ijrt}) \right],$$

which, given the *iid* extreme value varies over consumers and stores, reduces to:

$$CW_t = \frac{1}{\alpha} \ln \left( 1 + \sum_{j=1}^{J_t} \exp(\delta_{jt}^r) \right) + K_{1t}, \quad (21)$$

where  $K_{1t}$  is an unknown constant.

2. **Nested logit:** This computation is an immediate extension of the logit derivation discussed in equation (21) above:

$$CW_t = \frac{1}{\alpha} \ln \left\{ 1 + \sum_{g=1}^{G_t} \left[ \sum_{l \in g} \exp \left( \frac{\delta_{lt}^r}{\sigma_g} \right) \right]^{\sigma_g} \right\} + K_{2t},$$

where the term in curly brackets corresponds to the natural log of the denominator of  $s_{gt}$  formulated in equation (8).

3. **Random coefficients nested logit:** Here, we readapt a formula from Nevo (2000b) and write,

$$W_t = \frac{1}{ns} \sum_{i=1}^{ns} \frac{1}{\alpha_i} \left\{ \sum_{g=1}^{G_t} \left[ \sum_{l \in g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} v_i^p}{\sigma_g} \right) \right]^{\sigma_g} \right\} + K_{3t},$$

where  $\alpha_i$  is the random coefficient introduced in equation (11).

Note that in none of the above specifications we are able to compute the level of expected consumer surplus: only changes in expected consumers surpluses can be identified. Consequently, in our welfare computations we will choose a specification as reference category, and compute the changes in welfare generated by deviations from that default scenario.

### 3.4 Counterfactual

We have seen that our data on store coverage by products suggest the existence of “congestion” effects. A natural counterfactual would impose standard logit errors with a “standard logit model”. We do this by assuming all products have full coverage ( $r_{jt} = r_0 = 1$ , and  $\sum_{l=0}^{J_t} r_{lt} = J_t + 1, \forall t$ ). In essence, stores are asked to create (free of charge and without affecting inventories) extra space in their shelves to allow for the presence of all CSD products.

We keep as primitives the parameters of demand and marginal cost functions that we have estimated observing our true measure  $r_{jt}$  of store coverage, and under the assumption of full retail coverage, we compute the new equilibrium price in the market. In order to calculate the new equilibrium we use *fsolve*

in Matlab and unravel, for every period, a system of  $J_t$  nonlinear equations in  $\mathbf{p}_t$  (the pricing equations with the demand equations substituted in). The procedure for each of the three demand specifications is reported in appendix 6.B.<sup>14</sup> Having solved for the new equilibrium, we recompute the price elasticities at the new prices and market shares. We also calculate the change in welfare that results from moving from one equilibrium to another.<sup>15</sup>

## 4 Econometrics and Results

We need to estimate the demand and “supply” equations derived in the previous section. An important issue in empirical IO is how to deal with the endogeneity in pricing. Prices are endogenous due to their correlation with demand unobservables (through unobserved quality) and supply unobservables (through the markup function). The empirical IO literature has well documented price endogeneity due to the fact that part of a given product’s quality is observed by consumers and producers, but not by the econometrician. In the simple logit specification this translates into an upward bias of the Least Squares estimation of the price parameter (due to an underlying positive correlation between price and quality).<sup>16</sup> By-products of this bias include: inelastic estimates of demand elasticities of substitution, overestimation of markups and, in extreme cases, negative marginal costs. The literature has however suggested useful instruments to overcome this issue [Hausman and Taylor (1981), Berry (1994), BLP, Fershtman et al. (1999), Nevo (2000a, 2000b, 2001), amongst others].

In our paper, on top of this simultaneity issue, we also consider the role of an important unobservable variable in consumer taste. Product coverage of stores is a typical (relevant) omitted variable in the product-level estimation literature.<sup>17</sup> As an omitted variable, depending on the underlying correlation, it could worsen, improve, or leave unaltered the endogeneity effect. In appendix 6.C we delineate a theoretical aside on how simultaneity and omitted variable effects can interact. Regardless, observing store coverage may not be the panacea

<sup>14</sup>The same appendix also provides some useful algebra on the three demand specifications and their counterfactual.

<sup>15</sup>We rely here on the assumption of uniqueness of equilibria, but we are aware that the nonlinearity of the  $J_t$  system of equations is a fertile soil for multiple equilibria. If this is the case, we hope the equilibrium we pick up is the only reasonable one in economic terms.

<sup>16</sup>Note that we expect the price parameter to be negative, and thus the upward bias in our case would translate into a lower absolute value of the associated parameter.

<sup>17</sup>It is fair to add that this latter bias could be eliminated if one could find “ideal instruments” that simultaneously: i) mitigate (or annul) the simultaneity bias induced by unobserved product quality and ii) yield results that are uncorrelated with store coverage (see appendix 6.C for details).



to the problem, as store coverage could be correlated with unobserved product quality, and therefore be itself an endogenous variable (something that we will test in our econometrics). The above discussion can be extended to the nested logit specification, although in this case, the sign of the price parameter bias is less intuitive. The way in which the within segment market share relates to both unobserved quality and to price will be important.

We address the problem of endogeneity using a set of instruments, and in keeping with BLP, we postulate the following conditional moments independencies,

$$\begin{aligned} E(\xi_{jt}|\mathbf{z}_{1t}) &= E(\omega_{jt}|\mathbf{z}_{2t}) = 0 \\ E[(\xi_{jt}, \omega_{jt})'(\xi_{jt}, \omega_{jt})|\mathbf{z}_t] &= \Omega(\mathbf{z}_{jt}), \end{aligned} \quad (22)$$

where  $\mathbf{z}_t \equiv \{\mathbf{z}_{1t}, \mathbf{z}_{2t}\}$ . The second condition in (22) allows for heteroskedasticity in demand and supply unobservables' conditional variances.

We separate the parameters in a linear set  $\boldsymbol{\theta}_1$ , and a non-linear set  $\boldsymbol{\theta}_2$ , with  $\boldsymbol{\theta} \equiv \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\}$ . We combine linear and nonlinear GMM estimators, as in BLP and Nevo (2000a). As part of our estimation procedure, we compute an estimator of the asymptotic variance-covariance of the linear and non-linear parameters. This requires us to calculate the gradient of the GMM functions with respect to the parameters and, whenever possible, we make use of analytical solutions. We succeed in providing analytical solutions for the gradient of each demand side specification, and employ numerical derivations for the more complex "supply" side. The analytical gradient of the nonlinear parameters entering the demand side random coefficient nested logit requires a rather tedious derivation, which we outline in Appendix 6.D.

## 4.1 Instruments

A heated debate among those that use the GMM estimator is the choice of instruments. Newey (1990) shows the efficient set of instruments for nonlinear models to be equal to the conditional expectation of the derivative of the conditional moments with respect to the parameters vector. Unfortunately this conditional expectation is very difficult, if not impossible, to compute utilizing our nonlinear estimation. Consequently, we resort to selecting our instruments based upon their power and validity in conjunction with their underlying economic significance. Appendix 6.E fully details our choice of the instruments, which relies on a four step procedure. In step one we investigate the power of

the instruments by regressing our endogenous variables over alternative sets of instruments. In step two, we select the significant instruments and compute a demand side GMM to verify the validity of the chosen instruments. Step three avails of the estimated demand primitives from step two and deduces estimated marginal costs; accessing the power of supply side variables. The last step, checks for the validity of the instruments for those endogenous variables that directly, or via functional forms, enter the GMM pricing equation. The idea is to explore the power and validity of the instruments separately for demand and supply, which will help in our choice of instruments for the joint estimates of demand and supply.

Below we list and define the instruments that performed best in the correction of the price endogeneity and, where applicable, the store coverage endogeneity. These include BLP type instruments, such as the number of products and the sum of the product characteristics that proxy for the intensity of competition or product differentiation in our structural model of equilibrium. We divide these two types of instruments into summations over products belonging to the same firm and summations over products of other firms, for each period, as follows:

- $z_1$ : the sum of other products belonging to the firm;
- $z_2$ : the sum of months to stock out of other products of the firm;
- $z_3$ : the sum of log store coverage of other products of the firm (when store coverage is observable and exogenous);
- $z_4$ : the sum of products of other firms;
- $z_5$ : the sum of months to stock out of products of other firms;
- $z_6$ : the sum of log store coverage of products of other firms (when store coverage is observable and exogenous).

In addition we use Hausman and Taylor (1981) type instruments, aimed at capturing common underlying cost shifters:

- $z_7$ : initial condition of average price of products outside the segment the firm belongs to;
- $z_8$ : initial condition of average log store coverage of products outside the segment the firm belongs to (when store coverage is observable and endogenous).

## 4.2 Results

In each of our specifications we explore the following scenarios in which: i) store coverage is not observable; ii) store coverage,  $sc$ , is proxied by the  $AR$  variable (predicted from the nonlinear estimates set out in Table 3); iii) store coverage,

is observable.<sup>18</sup>

The Least Squares estimates of our demand systems are reported in Table 4. If we look at the  $R^2$  for our logit model we see that when store coverage is either omitted, or proxied for by  $AR$ , only 36% of the variability of demand is explained. On the other hand, once we account for store coverage with actual data, a good 83% of the variability is captured. The nested logit specification has the advantage of reducing the discrepancy among the three scenarios, by moving from an  $R^2$  of almost 88%, in case of omitted or proxied store coverage, to an  $R^2$  over 91% in case of store coverage being observable. The table also provides evidence that the  $AR$  proxy helps little in correcting for the bias that the omitted coverage produces in the price and log of within segment market share coefficients. Our observable store coverage variable is much more informative, being both product- and period-specific. For example, if we compare Columns 1 and 3 for the logit case, we observe the estimated price coefficient to be (in absolute value) lower, when we account for store coverage. This direction of the estimated price effect suggests that the simultaneity bias is amplified when one allows for store coverage. However, such a tendency in the price parameter is not shared by the nested logit specifications. As we can see from a comparison of columns 4 and 6, the price parameter here rises (in absolute value). This change in direction is caused by the correlation that the  $\log$  of within market shares and  $\log$  of store coverage have with prices (as one can note from Table 5).

A common feature of the various scenarios and specifications of Table 4 is the high number of estimated negative marginal costs, which we have discussed as being driven primarily by a simultaneity bias. We tackle the simultaneity bias by jointly estimating the demand and “supply” equations by GMM. As already anticipated in the previous section, the selection of instruments relies upon a trade-off between power and validity. We present descriptive statistics of the instruments and their correlations, along with some relevant variables, in Tables 5 and 6. We report our final choice of instruments in a footnote of Table 7.

If we compare the Least Squares estimates for the logit and nested logit specifications of Table 4, with the equivalent joint GMM estimates of Table 7, we note that both the  $(-\alpha)$  price parameter, and the  $\sigma_g$  within segment market share, are heavily downward biased in the OLS estimations. On the other hand, the store coverage parameter  $\sigma_r$  is only downward biased in the nested logit

---

<sup>18</sup>Due to its poor performance, we do not include the  $AR$  variable in the random coefficient nested logit specification.

specification that has observable store coverage.

The demand side of Table 7 separately reports the linear and nonlinear estimated parameters.<sup>19</sup> The store coverage parameter  $\sigma_r$ , is not estimated well in the logit specifications, as the estimate exceeds the unitary theoretical upper bound of its support. Only under the richer nested logit specification do we get reasonable values for  $\sigma_r$ . We then observe that using the actual store coverage in the nested logit and random coefficient nested logit specifications brings about the estimate of  $\sigma_r$  to a value around 0.7, as it corrects for the bias.

We now turn our attention to the nonlinear parameters. The negative of the mean price effect estimated coefficient ( $-\alpha$ ) is significant and varies between 3 and 3.8 among the different specifications. We do not find the standard deviation parameter of the random coefficient nested logit specification  $\sigma_p$ , to be significantly different from zero, particularly when store coverage is observable. This is due to a lack of individual observable characteristics in our data, and subsequently to a poor set of available instruments. Another nonlinear parameter is the one attached to the within segment market share  $\sigma_g$ . Only the nested logit specification based on the *AR* proxy poorly estimates  $\sigma_g$ . The other specifications and scenarios furnish a significant estimated parameter in the range of 0.55 and 0.65.

The results on the supply side do not provide evidence of a significant role of inventory and store coverage in the marginal cost function for the logit or nested logit specifications, but are significant and of the right sign in the random coefficient models. Declining inventories increase cost, and store coverage reduces cost. Still, firm, packaging and seasonal dummies do capture most of the marginal cost variability.

Table 7 also displays useful statistics that can help us select the best specification. The pseudo  $R^2$  of the demand side  $R_D^2$ , emphasizes the strong role that the observability of store coverage has in explaining the pure product-level variability. As for the supply side, the pseudo  $R_S^2$  is always above 70 percent. The  $J$ -statistic is informative on the validity of the instruments. Figures of the p-value below 0.05 indicate that some of the moment conditions are not supported by the data at the 5 percent level. An important statistic is the number of negative marginal costs. We elaborate on this outcome for counterfactuals when we use a selected demand model.

Based on the above discussion on the estimated parameters and various statistics, we feel confident that the best estimation results come from the nested logit with observed store coverage. This one shows all but one marginal cost to

---

<sup>19</sup>Nonlinear parameters are those that enter the markup function.

be nonnegative; a good overidentification test on the instruments to accept the null hypothesis of valid moment restrictions; and the average estimated part of the product-level variability for demand and supply to be around 80 percent. In addition, the estimated parameters all seem reasonable from an economic point of view.

Table 8 presents the estimates of the average price elasticity of substitution, price and marginal cost, over the last period  $T$ .<sup>20</sup> Our calculation of the estimates of own and cross price elasticities at the product level, are based on the formulae developed in appendix 6B. We aggregate (market share weighted) over primitives of products by defined groupings in terms of store coverage (low and high; full in the counterfactual). We have 90 products with a coverage below 50 per cent, and 64 products with a coverage above 50 per cent coverage of the market. Within the groupings, the cross price elasticities that are presented are a weighted sum of cross price effects for each product. The computed price per liter is also a weighted sum, as is the estimated marginal cost per product. We do our analysis using estimates based on parameters from the Logit and Nested Logit demand systems to demonstrate our results. We wish to compare our demand and supply outcomes in the case when we use the actual data on store coverage, to outcomes produced from our counterfactual that imposes full store coverage of all products in the market. This gives us an idea of what the outcomes would look like if consumers did not suffer disutility from products being unavailable in stores, which is the assumption embedded in standard logit errors. We express, in millions of Irish pounds, the change in profit and consumer welfare that result from moving to full coverage. This is the result of the difference between outcomes in the counterfactual and those in the preferred scenario where store coverage observable.

In Table 8 we first analyze the logit model. The results of the counterfactual indicate that products starting from low coverage do have a jump in the intensity of competition in the market coming from cross price effects. In addition, own prices and costs do go up. In products starting from a position of high coverage, while the cross-price effects go up, the own price and costs go down. Overall, profits in the market increase as does consumer welfare. This is because of the increased competition coming from the smaller products expanding coverage steals market share in both the inside and outside markets. Once again, in the nested logit model we compare our results using actual store coverage data to that simulated when full store coverage by each product is imposed on the

---

<sup>20</sup>For brevity, we show only the last period, although we have computed elasticities, price equilibrium, and marginal costs for all periods.

model. The results are similar to the logit model. In products starting from a low coverage we see a jump in the intensity of cross price competition in the market with own prices and costs rising. In products starting from a position of high coverage we also see a jump in the intensity of cross price competition but own price and costs do go down. Overall profits in the market increase as does consumer welfare. This welfare gain is weaker than the logit model, as the segmentation of the market protects products from increased competition from those outside their segment or from the outside good.

In Table 9 we take the own and cross price effects of all products in every period within ten intervals of store coverage and test whether the estimated price effects in the estimated logit and nested logit models with full coverage are different to the counterfactual that imposes full coverage in each of these ten intervals. We present the results of the Kolmogorov-Smirnov test for equality of distributions. We find that while own price effects are not so different, the distribution of cross price effects are radically different. The own price effect is an outcome of two forces. For low coverage products the counterfactual reduces disutility for the product greatly, leading to an expansion in market share and downward pricing. High coverage products benefit less from the counterfactual. They lose market share and charge lower prices to protect profits. Overall, own price effects do not change too much but their components do.

Tables 10 repeats the analysis of Table 8, but this time we focus on demand and cost outcomes within the bigger segments of the market. In particular, Cola, Orange, Lemonade and Mixed Fruit flavour segments in Standard and 2-Liter packaging. The general result is that using standard logit errors would lead to overestimates of the demand (welfare and price elasticities) and cost outcomes. Ignoring product congestion inside stores by using standard logit errors can significantly bias estimates of price elasticities and costs upwards. Markets would be estimated to be far more competitive than they really are. Any counterfactual that is based on these primitives is also likely to generate inaccurate results. In summary, our results support the Monte Carlo results of *AR*.

## 5 Conclusions

Consumers can face a congestion in their choice set of products within stores. Following Akerberg and Rysman (2005) we allow logit errors to represent idiosyncratic unobserved consumer preferences over retail stores and products. Having product level data on store coverage we estimate their logit, nested logit

and random coefficients logit models of product demand jointly with cost in a structural model of equilibrium for products in retail Carbonated Soft Drinks. Allowing for store coverage has a very significant impact on the estimated parameters and the predictive power of our structural model. In our counterfactual we impose full-coverage (no congestion) on our data and evaluate demand price elasticities and welfare. We see that the own and cross price elasticities generally get larger in the new equilibrium. Building structural models of industries with standard logit errors overlooks a key aspect of product differentiation, i.e. consumers having different choice sets inside stores. Products inside markets tend to be more protected from competition than allowed by the modeler. As we show this leads to inaccurate oversized estimates of demand price elasticities and consumer welfare inside industries. Any subsequent counterfactual on the likely effect of mergers or price coordination could lead to very misleading conclusions.

## References

- [1] Akerberg, D., Benkard, L., Berry, S. and A. Pakes (2007), “Econometric Tools for Analyzing Market Outcomes” in *Handbook of Econometrics*, Vol. 6A, J.J. Heckman and Leamer E. (eds.), North Holland, 4171-276.
- [2] Akerberg, D. and M. Rysman (2005), “Unobserved Product Differentiation in Discrete-Choice Models: Estimating Price Elasticities and Welfare Effects,” *The RAND Journal of Economics*, 36, 771-88.
- [3] Arcidiacono, P. (2005), “Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?,” *Econometrica*, 73, 1477-524.
- [4] Bajari, P. and C. L. Benkard (2005), “Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach,” *Journal of Political Economy*, 113, 1239-76.
- [5] Ben-Akiva, M. and S. Lerman (1985), “Discrete Choice Analysis: Theory and Application to Travel Demand,” MIT Press, Cambridge, MA.
- [6] Berry, S.T. (1994), “Estimating Discrete Choice Models of Product Differentiation,” *The RAND Journal of Economics*, 25, 242-62.
- [7] Berry, S., Levinsohn, J. and A. Pakes (1995), “Automobile Prices in Market Equilibrium,” *Econometrica*, 63, 841-90.

- [8] Berry, S.T. and A. Pakes (2007), "The Pure Characteristics Demand Model," *International Economic Review*, 48, 1193-225.
- [9] Berry, S.T and J. Waldfogel (1999), "Free Entry and Social Inefficiency in Radio Broadcasting," *The RAND Journal of Economics*, 30, 397-420.
- [10] Bresnahan, T.F, Stern, S., and M. Trajtenberg (1997), "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *The RAND Journal of Economics*, 28, S17-44.
- [11] Cardell, N.S. (1997), "Variance Components Structures for the Extreme Value and Logistic Distributions with Applications to Models of Heterogeneity," *Econometric Theory*, 13, 185-213.
- [12] Crawford, G. (2000), "The Impact of the 1992 Cable Act on Household Demand and Welfare," *The RAND Journal of Economics*, Vol. 31(2000), pp. 422-49.
- [13] Crawford, G. and M. Shum (2005), "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73, 1137-73.
- [14] D'Aspremont, C., Gabszewicz, J. and J. Thisse (1979), "On Hotelling's Stability in Competition," *Econometrica*, 47, 1145-50.
- [15] Davis, P. (2006), "Spatial Competition in Retail Markets: Movie Theatres," *The RAND Journal of Economics*, 37, 964-82.
- [16] Deaton A. and G. Laroque (1992), "On the Behaviour of Commodity Prices," *Review of Economic Studies*, 59, 1-23.
- [17] Deaton A. and G. Laroque (1996), "Competitive Storage and Commodity Price Dynamics", *The Journal of Political Economy*, 104, 896-923.
- [18] Fershtman C., Gandal N., and Markovich S., (1999). "Estimating the Effect of Tax Reform in Differentiated Product Oligopolistic Markets," *Journal of Public Economics*, 151-70, Vol. 74.
- [19] Hausman, J. and W. Taylor (1981), "Panel Data and Unobservable Individual Effects," *Econometrica*, 49, 1377-98.
- [20] Hotelling, H. (1929), "Stability in Competition," *Economic Journal*, 39, 41-57.



- [21] Mariuzzo, F. Walsh, P. P. and C. Whelan (2003), "Firm Size and Market Power in Carbonated soft Drinks," *Review of Industrial Organization*, 23, 283-99.
- [22] Nevo, A. (2000a), "A Practitioners Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9, 513-48.
- [23] Nevo, A. (2000b), "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *The RAND Journal of Economics*, 31, 395-421.
- [24] Nevo, A. (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307-42.
- [25] Newey, W. K. (1990), "Efficient Instrumental Variables Estimation of Non-linear Models," *Econometrica*, 58, 809-37.
- [26] Pakes, A. (2003), "A Reconsideration of Hedonic Price Indexes with an Application to PCs," *American Economic Review*, 93, 1578-93.
- [27] Perloff, J. M., Karp L. S., and A. Golan (2007), *Estimating Market Power and Strategies*, Cambridge, MA: MIT Press.
- [28] Petrin, A. (2002), "Quantifying the Benefits of New Products: the Case of the Minivan," *Journal of Political Economy*, 110, 705-29.
- [29] Rysman, M. (2004), "Competition Between Networks: A Study of the Market for Yellow Pages," *Review of Economic Studies*, 71, 483-512.
- [30] Salop, S. (1979), "Monopolistic Competition with Outside Goods", *The Bell Journal of Economics*, 10, 141-56.
- [31] Small, K. A. and S. Rosen (1981), "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, 49, 105-30.
- [32] Song, M. (2007), "Measuring Consumer Welfare in the CPU Market: An Application of the Pure Characteristics Demand Model," *The RAND Journal of Economics*, 38 ,429-46.
- [33] Sutton, J. (1991), *Sunk Costs and Market Structure: Price Competition, Advertising and the Evolution of Concentration*, Cambridge, MA: MIT Press.

- [34] Sutton, J. (1998), *Technology and Market Structure: Theory and History*, Cambridge, MA: MIT Press.
- [35] Town, R. and S. Liu (2003), “The Welfare Impact of Medicare HMOs,” *The RAND Journal of Economics*, 34, 719-36.
- [36] Train, K. E. (2003), *Discrete Choice Methods with Simulations*, Cambridge University Press.
- [37] Walsh, P.P. and C. Whelan (1999), “Modelling Price Dispersion as an Outcome of Competition in the Irish Grocery Market,” *Journal of Industrial Economics*, 47, 325-43.

**Table 1:** Segments quantity and price levels, averaged Oct.'92 – May '97

	Products	Firms	Price Per Liter	Unit Sales Ltr (000)	Unit Sales Shares	Revenue £IR(000)	Revenue Shares
Cola Can	6	5	1.43	1,486	4.22	2,116	8.02
Cola Standard	11	5	1.26	1,333	3.78	1,692	6.41
Cola 1.5 liter	3	3	0.75	893	2.53	672	2.55
Cola 2 liter	5	4	0.50	3,867	10.97	1,945	7.37
Cola Can Multipacks	5	2	0.96	678	1.92	648	2.46
Orange Can	6	4	1.38	653	1.85	887	3.36
Orange Standard	10	6	1.27	741	2.10	931	3.53
Orange 1.5 liter	5	4	0.68	781	2.22	535	2.03
Orange 2 liter	5	4	0.46	3,000	8.51	1,382	5.24
Orange Can Multipacks	3	3	0.97	174	0.49	170	0.64
Lemonade Can	4	2	1.41	498	1.41	699	2.65
Lemonade Standard	5	2	1.16	487	1.38	569	2.16
Lemonade 1.5 liter	3	2	0.71	1,323	3.75	939	3.56
Lemonade 2 liter	4	2	0.47	4,140	11.75	1,941	7.35
Lemonade Can Multipacks	2	1	0.97	128	0.36	124	0.47
Mixed Fruit Can	7	5	1.39	752	2.13	1,045	3.96
Mixed Fruit Standard	19	10	1.37	2,217	6.29	3,128	11.85
Mixed Fruit 1.5 liter	7	6	0.74	633	1.80	465	1.76
Mixed Fruit 2 liter	8	6	0.41	6,612	18.76	2,635	9.99
Mixed Fruit Can Multipacks	1	1	0.83	8	0.02	6	0.02
Diet Cola Can	4	3	1.39	392	1.11	542	2.05
Diet Cola Standard	3	3	1.30	328	0.93	424	1.61
Diet Cola 1.5 liter	4	2	0.75	293	0.83	221	0.84
Diet Cola 2 liter	4	3	0.55	1,005	2.85	537	2.03
Diet Cola Can Multipacks	3	2	0.96	222	0.63	213	0.81
Diet Orange Can	2	1	1.27	83	0.23	106	0.40
Diet Orange Standard	1	1	1.19	16	0.05	19	0.07
Diet Orange 1.5 liter	1	1	0.71	76	0.21	54	0.20
Diet Orange 2 liter	3	2	0.56	254	0.72	141	0.53
Diet Lemonade Can	2	2	1.44	186	0.53	268	1.01
Diet Lemonade Standard	1	1	1.29	75	0.21	96	0.36
Diet Lemonade 1.5 liter	1	1	0.73	572	1.62	415	1.57
Diet Lemonade 2 liter	2	1	0.59	1,198	3.40	699	2.65
Diet Lemonade Can Multipacks	1	1	0.96	74	0.21	71	0.27
Diet Mixed Fruit Can	2	2	1.27	14	0.04	18	0.07
Diet Mixed Fruit Standard	2	2	1.17	14	0.04	17	0.06
Diet Mixed Fruit 1.5 liter	1	1	0.83	1	0.00	1	0.00
Diet Mixed Fruit 2 liter	1	1	0.55	40	0.11	22	0.08
<b>Total</b>	<b>157</b>	<b>107</b>		<b>35,249</b>	<b>100</b>	<b>26,388</b>	<b>100</b>

**Table 2:** Segments store coverage and inventory levels, May '93 (initial), May '95 (middle) and May '97 (end)

	Coverage Initial	Coverage Middle	Coverage End	Inventories Initial	Inventories Middle	Inventories End
Cola Can	0.95	0.93	0.94	0.53	0.53	0.67
Cola Standard	0.69	0.86	0.89	0.60	0.47	0.53
Cola 1.5 liter	0.84	0.84	0.81	0.70	0.53	0.63
Cola 2 liter	0.68	0.74	0.79	0.47	0.40	0.43
Cola Can Multipacks	0.43	0.49	0.50	0.53	0.37	0.37
Orange Can	0.80	0.80	0.75	0.83	0.70	0.90
Orange Standard	0.81	0.72	0.71	0.83	0.53	0.67
Orange 1.5 liter	0.64	0.65	0.61	0.73	0.57	0.77
Orange 2 liter	0.70	0.61	0.62	0.43	0.43	0.53
Orange Can Multipacks	0.37	0.39	0.43	0.63	0.63	0.47
Lemonade Can	0.91	0.95	0.84	0.87	0.60	0.90
Lemonade Standard	0.81	0.74	0.82	0.87	0.60	0.60
Lemonade 1.5 liter	0.83	0.87	0.87	0.57	0.37	0.50
Lemonade 2 liter	0.76	0.75	0.80	0.50	0.40	0.37
Lemonade Can Multipacks	0.38	0.35	0.40	0.63	0.33	0.53
Mixed Fruit Can	0.87	0.83	0.77	0.80	0.60	0.93
Mixed Fruit Standard	0.92	0.80	0.76	0.80	0.57	0.73
Mixed Fruit 1.5 liter	0.54	0.47	0.41	0.97	0.60	0.87
Mixed Fruit 2 liter	0.68	0.65	0.64	0.60	0.33	0.50
Mixed Fruit Can Multipacks	0.84	0.80	0.04	0.70	0.63	0.10
Diet Cola Can	0.78	0.82	0.88	0.80	0.67	0.77
Diet Cola Standard	0.55	0.76	0.89	0.50	0.77	0.63
Diet Cola 1.5 liter	0.40	0.60	0.75	0.47	0.43	0.70
Diet Cola 2 liter	0.69	0.41	0.68	1.00	0.33	0.43
Diet Cola Can Multipacks	0.51	0.69	0.41	0.93	0.70	0.27
Diet Orange Can	0.45	0.46	0.70	0.53	0.77	0.90
Diet Orange Standard	0.84	0.48	0.52	0.80	0.40	0.73
Diet Orange 1.5 liter	0.84	0.91	0.35	0.67	0.53	1.13
Diet Orange 2 liter	0.60	0.54	0.51	0.43	0.43	0.50
Diet Lemonade Can	0.35	0.85	0.90	0.70	0.50	0.80
Diet Lemonade Standard	0.49	0.73	0.75	0.87	0.33	0.23
Diet Lemonade 1.5 liter	0.01	0.32	0.84	0.10	0.27	0.53
Diet Lemonade 2 liter	0.36	0.22	0.77	0.67	1.30	0.33
Diet Lemonade Can Multipacks	0.95	0.33	0.35	0.53	0.23	0.43
Diet Mixed Fruit Can	0.69	0.93	0.17	0.60	0.53	1.23
Diet Mixed Fruit Standard	0.84	0.86	0.37	0.70	0.47	0.67
Diet Mixed Fruit 1.5 liter	0.68	0.84	0.38	0.47	0.53	0.40
Diet Mixed Fruit 2 liter	0.43	0.74	0.94	0.53	0.40	0.67
<b>Average</b>	<b>0.66</b>	<b>0.68</b>	<b>0.65</b>	<b>0.66</b>	<b>0.52</b>	<b>0.61</b>

Table 3: Nonlinear Least Squares estimations of retail store coverage

Parameters	Specification (i)	Specification (ii)
$t$	0.020*** (0.007)	0.015*** (0.009)
$\frac{t^2}{100}$	-0.044*** (0.017)	-0.035*** (0.019)
Cons	3.954*** (0.108)	3.498*** (0.794)
$\lambda_1$	1.000*** (0.001)	
$\lambda_2$		-0.900*** (0.179)
<b>STATISTICS</b>		
$R^2$	0.92	0.92
$N$	28	28
<b>TESTS</b>		
$\lambda_1 = 1$	-0.000 (0.001)	
$\lambda_2 = -1$		0.100 (0.179)

Standard errors in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 4: OLS demand estimation. Dependent variable  $\ln(s/s_0)$ 

Parameters	OLS	OLS AR	OLS SC	NOLS	NOLS AR	NOLS SC
Logit			Nested Logit			
$\beta_1$ [cons]	-5.619*** (0.503)	-4.236*** (0.645)	-5.473*** (0.258)	-2.876*** (0.221)	-2.639*** (0.284)	-3.759*** (0.186)
$\beta_2$ [inv]	-21.156*** (1.015)	-21.192*** (1.013)	-5.866*** (0.543)	-4.224*** (0.466)	-4.240*** (0.466)	-3.071*** (0.390)
$-\alpha$ [p]	0.925*** (0.159)	0.923*** (0.158)	0.746*** (0.081)	0.029 (0.070)	0.029 (0.070)	0.257*** (0.059)
$\sigma^r$ [ $\Gamma^\dagger$ ]		1.869*** (0.548)	1.091*** (0.011)		0.323 (0.241)	0.501*** (0.013)
$(1 - \sigma^g)$ [ $l_{s/g}$ ]				0.912*** (0.008)	0.911*** (0.008)	0.596*** (0.010)
<b>STATISTICS</b>						
N.	3644	3644	3644	3644	3644	3644
N. mc<0	2733	2741	3317	3644	3644	3525
$R^2$	0.360	0.363	0.832	0.877	0.877	0.914

Standard errors in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1,  $\dagger$  AR proxy in the AR scenarios. Not reported firm, segment, and (package x time) fixed effects.

Table 5: Instruments and main variables correlation matrix

	s	p	lsg	stkout	lrj	z1	z2	z3	z4	z5	z6	z7	z8
s	1.00												
p	-0.31	1.00											
lsg	0.39	-0.13	1.00										
inv	-0.18	0.10	-0.29	1.00									
lrj	0.32	-0.08	0.78	-0.30	1.00								
z1	0.05	0.01	0.24	-0.07	0.12	1.00							
z2	0.05	0.01	0.24	-0.06	0.14	0.94	1.00						
z3	-0.07	-0.02	-0.15	0.08	-0.04	-0.90	-0.81	1.00					
z4	-0.05	0.03	-0.33	0.04	-0.25	-0.65	-0.66	0.32	1.00				
z5	-0.05	0.04	-0.29	0.04	-0.23	-0.54	-0.55	0.23	0.90	1.00			
z6	0.04	-0.04	0.28	-0.02	0.24	0.32	0.35	0.01	-0.91	-0.85	1.00		
z7	0.03	-0.04	0.03	0.02	0.07	0.16	0.16	-0.10	-0.20	-0.18	0.16	1.00	
z8	-0.01	0.25	0.12	-0.04	0.06	-0.06	-0.07	0.01	0.11	0.10	-0.10	-0.62	1.00

Table 6: Descriptive statistics main variables and instruments

Variable	Mean	Std. Dev.	Min	Max
s	0.00	0.01	5.98E-06	0.08
p	0.99	0.43	0.29	4.54
lsg	-2.54	2.04	-8.94	0
inv	0.03	0.03	0.00	0.68
lrj	-1.33	1.57	-6.91	0
z1	32.65	17.97	0	57
z2	27.01	14.84	0	55.78
z3	-40.11	26.14	-94.79	0
z4	99.94	26.33	50	153
z5	86.93	23.86	40.37	147.28
z6	-142.92	65.32	-279.09	-39.81
z7	0.21	0.38	0	1.28
z8	-0.21	0.47	-2.86	0

Table 7: GMM simultaneous estimation of demand and supply

DEMAND (D)	Parameters	L	L AR	L SC	NL	NL AR	NL SC	RC	RC SC
Linear	$\beta_1$ [cons]	-4.290*** (1.176)	-3.366** (1.472)	-2.652*** (1.014)	-4.248*** (1.535)	-3.481*** (1.644)	-2.853*** (0.528)	-3.231*** (0.472)	-2.219*** (0.322)
	$\beta_2$ [inv]	-0.713*** (0.070)	-0.714*** (0.070)	-0.210*** (0.060)	-0.503*** (0.125)	-0.641*** (0.105)	-0.150*** (0.057)	-0.419*** (0.212)	-0.130*** (0.047)
	$\sigma^r$ [ $lr^\dagger$ ]		1.179*** (0.316)	1.085*** (0.020)		0.325*** (0.102)	0.772*** (0.128)		0.637*** (0.084)
Nonlinear	$-\alpha$ [p]	3.303*** (1.279)	3.263*** (1.341)	3.092*** (1.112)	2.879** (1.486)	3.499** (1.563)	3.013*** (0.564)	3.838*** (0.936)	3.763*** (0.838)
	$(1 - \sigma^g)$ [ $ls/g$ ]				0.353** (0.171)	0.111 (0.140)	0.320*** (0.129)	0.479*** (0.204)	0.457*** (0.060)
	$\sigma^p$ [ $p\nu_i^p$ ]							0.710 (0.942)	0.417 (1.235)
SUPPLY (S)									
Linear	$\gamma_1$ [cons]	0.584*** (0.141)	1.808 (1.914)	0.557*** (0.140)	0.643*** (0.193)	1.530*** (0.179)	0.560*** (0.071)	0.718*** (0.119)	0.667*** (0.130)
	$\gamma_2$ [inv]	-0.004 (0.008)	-0.004 (0.008)	-0.005 (0.010)	0.002 (0.009)	-0.002 (0.002)	0.001 (0.002)	0.019* (0.008)	0.002 (0.010)
	$\gamma_3$ [ $lr^\dagger$ ]		1.466 (2.275)	-0.003 (0.007)		0.558*** (0.039)	0.000 (0.002)		-0.032* (0.020)
STATISTICS									
Pseudo $R_D^2$		0.332	0.314	0.798	0.657	0.412	0.855	0.250	0.661
Pseudo $R_S^2$		0.823	0.824	0.824	0.729	0.809	0.763	0.770	0.784
J-stat		7.258	1.819	10.418	14.541	6.165	4.313	29.045	13.955
p-val J-stat		0.065	0.611	0.034	0.006	0.191	0.437	0.000	0.016
N.		3644	3644	3644	3644	3644	3644	3644	3644
N. mc < 0		64	82	133	3	0	1	85	84
N. mc < 0 cft				158			0		88

Standard errors in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. Not reported firm, segment, and (package x time) fixed effects.  $\dagger$  AR proxy for the AR scenarios. The standard errors are robust and account for potential correlation between demand and supply unobservables at the product level. We have not computed the part of the asymptotic variance-covariance, induced by simulation errors, thus assuming, as in Nevo (2000a, 2001), this to be a small and not influential component

Instruments:

[L: D =  $z_2, z_7$ ; "S" =  $z_1$ ], [L AR: D =  $z_2, z_7$ ; "S" =  $z_1$ ], [L SC: D =  $z_3, z_7$ ; "S" =  $z_4, z_5$ ].

[NL: D =  $z_1, z_2, z_7$ ; "S" =  $z_7$ ], [NL AR: D =  $z_1, z_2, z_7$ ; "S" =  $z_5$ ], [NL SC: D =  $z_1, z_3, z_7$ ; "S" =  $z_4, z_5$ ].

[RCNL: D =  $z_1, z_2, z_7$ ; "S" =  $z_4, z_5$ ], [RCNL SC: D =  $z_1, z_3, z_7$ ; "S" =  $z_4, z_5$ ].

Table 8: Estimates of price elasticities of substitution and marginal cost

Weighted Averages using Market Shares of Products, by Low and High Store Coverage	Store coverage	# Products	Own-Price Elasticity	Sum Cross-Price Elasticity	Price per liter	Estimated Marginal Cost per liter
Logit Model						
SC Logit	Low	90	-2.54	1.41	.82	.42
Counterfactual	Full	90	-2.71	1.81	.89	.49
SC Logit	High	64	-2.19	1.37	.72	.31
Counterfactual	Full	64	-2.09	1.81	.69	.27
No Congestion-Full	Change in Total Profit = 3.5			Change in Consumer Welfare= 10.4		
Nested Logit						
SC Nested Logit	Low	90	-3.38	2.37	.82	.50
Counterfactual	Full	90	-3.56	2.55	.90	.58
SC Nested Logit	High	64	-2.61	1.80	.72	.37
Counterfactual	Full	64	-2.52	2.18	.66	.32
No Congestion-Full	Change in Total Profit = 1.71			Change in Consumer Welfare = 7.7		

Table 9: Kolmogorov-Smirnov equality of distributions test (reference distribution observed store coverage)

Proportion Store Coverage	[0 , .1)	[.1 , .2)	[.2 , .3)	[.3 , .4)	[.4 , .5)	[.5 , .6)	[.6 , .7)	[.7 , .8)	[.8 , .9)	[.9 , 1]
<b>N</b>	662	192	211	370	502	445	333	311	318	300
<b>Logit</b>										
L Cft (OP)	1.00	1.00	1.00	0.86	0.75	0.50	0.62	0.64	0.85	0.40
L Cft (CP)	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
<b>Nested Logit</b>										
NL Cft (OP)	<b>0.02</b>	1.00	1.00	0.97	0.75	0.93	0.87	0.50	0.18	<b>0.00</b>
NL Cft (CP)	<b>0.00</b>	0.26	<b>0.01</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>

Reject if p<.05 (Bold values). OP=Own Price Elasticities; CP=Cross Price Elasticities



Table 10: Estimates of price elasticities of substitution and marginal cost by segment

<b>Packaging</b> ( <i>Weighted Averages Using Market Shares of Brands, by Low and High Store Coverage</i> )	<b>Store Coverage</b>	<b># Brands</b>	<b>Own-Price Elasticity</b>	<b>Sum Cross-Price Elasticity</b>	<b>Price per liter</b>	<b>Estimated Marginal Cost per liter</b>
<b>Standard Cola</b>						
SC Nested Logit	.02	7	-6.41	3.02	1.45	1.18
Counterfactual	Full	7	-4.59	3.14	0.86	0.48
SC Nested Logit	.90	4	-4.31	1.98	1.23	0.85
Counterfactual	Full	4	-5.11	2.71	1.21	0.85
<b>2 Liter Cola</b>						
SC Nested Logit	.05	1	-1.88	2.06	0.42	0.20
Counterfactual	Full	1	-1.87	2.29	0.43	0.20
SC Nested Logit	.79	4	-1.65	1.56	0.49	0.11
Counterfactual	Full	4	-1.67	1.93	0.47	0.11
<b>Standard Orange</b>						
SC Nested Logit	.12	6	-5.38	3.03	1.22	0.96
Counterfactual	Full	6	-4.48	3.51	1.19	0.90
SC Nested Logit	.74	4	-4.80	2.61	1.19	0.91
Counterfactual	Full	4	-5.03	3.18	1.17	0.91
<b>2 Liter Orange</b>						
SC Nested Logit	.13	1	-2.36	2.02	0.54	0.20
Counterfactual	Full	1	-2.30	2.21	0.55	0.20
SC Nested Logit	.64	4	-1.89	1.89	0.47	0.19
Counterfactual	Full	4	-1.92	2.18	0.49	0.19
<b>Standard Lemonade</b>						
SC Nested Logit	.08	2	-5.61	2.98	1.38	0.88
Counterfactual	Full	2	-5.81	3.02	1.39	0.99
SC Nested Logit	.84	3	-4.37	2.24	1.17	0.81
Counterfactual	Full	3	-4.62	2.93	1.15	0.81
<b>2 Liter Lemonade</b>						
SC Nested Logit	.04	1	-2.23	2.01	0.50	0.10
Counterfactual	Full	1	-2.15	2.29	0.50	0.10
SC Nested Logit	.80	3	-1.69	1.58	0.49	0.12
Counterfactual	Full	3	-1.71	1.94	0.47	0.12
<b>Standard Mixed Fruit</b>						
SC Nested Logit	.28	10	-3.80	3.11	0.88	0.63
Counterfactual	Full	10	-4.73	3.13	1.13	0.88
SC Nested Logit	.87	9	-6.07	2.93	1.45	1.19
Counterfactual	Full	9	-6.10	3.26	1.40	1.15
<b>2 Liter Mixed Fruit</b>						
SC Nested Logit	.31	2	-1.91	1.81	0.43	0.16
Counterfactual	Full	2	-1.81	2.18	0.41	0.15
SC Nested Logit	.65	6	-1.52	1.45	0.38	0.11
Counterfactual	Full	6	-1.55	2.06	0.38	0.11

Figure 1: Bimonthly unit sales over the 28 bimonthly periods (JJ 93 is sales in June/July in 1993)

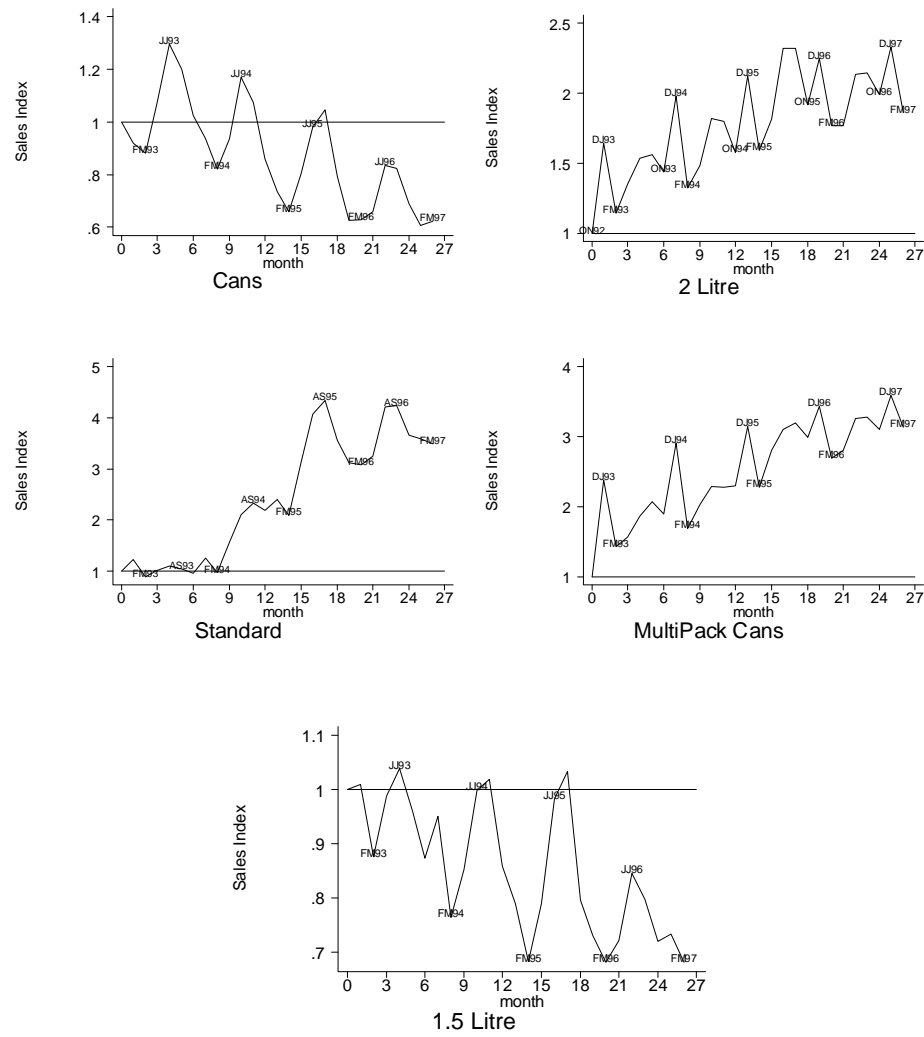


Figure 2: Actual store coverage (top panel), number of products (center panel), and estimated store coverage (bottom panel)

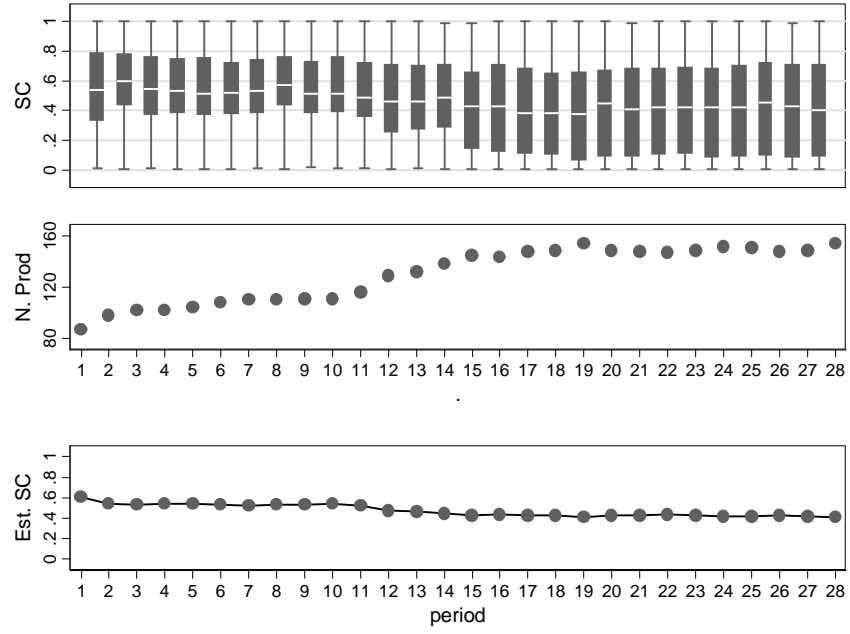
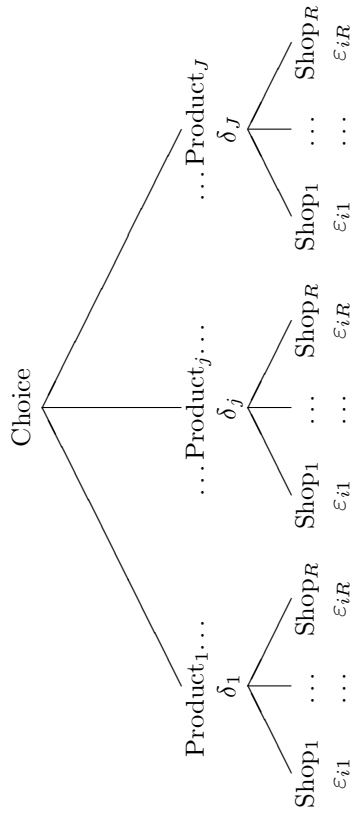


Figure 3: Classical Logit Model

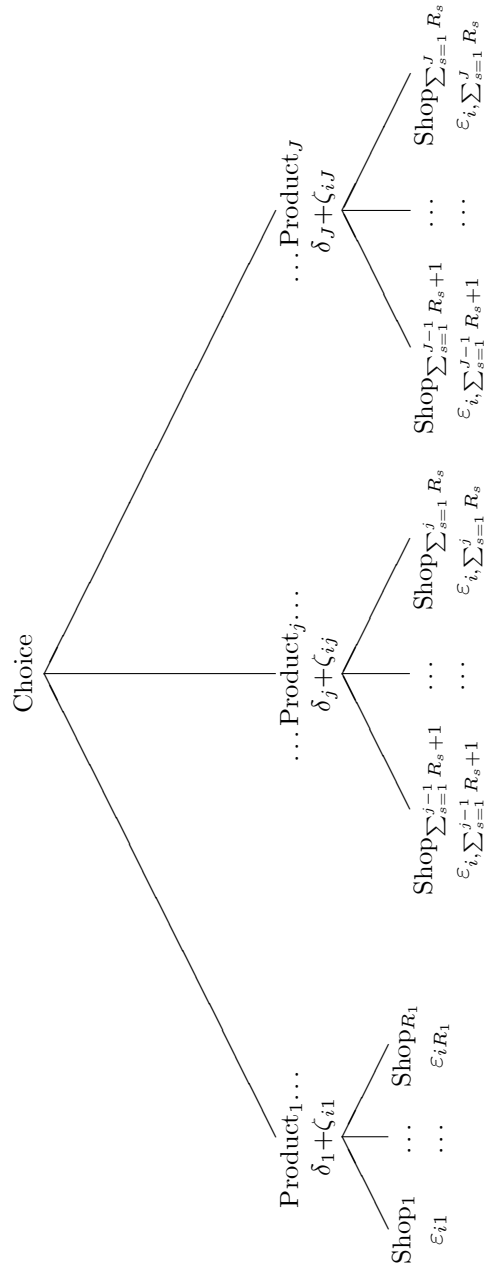
$$u_{ijr} = \delta_j + \varepsilon_{ir}$$



$$s_j = \frac{\exp(\delta_j)}{\sum_{l=1}^J \exp(\delta_l)}$$

Figure 4: Structural Logit Model

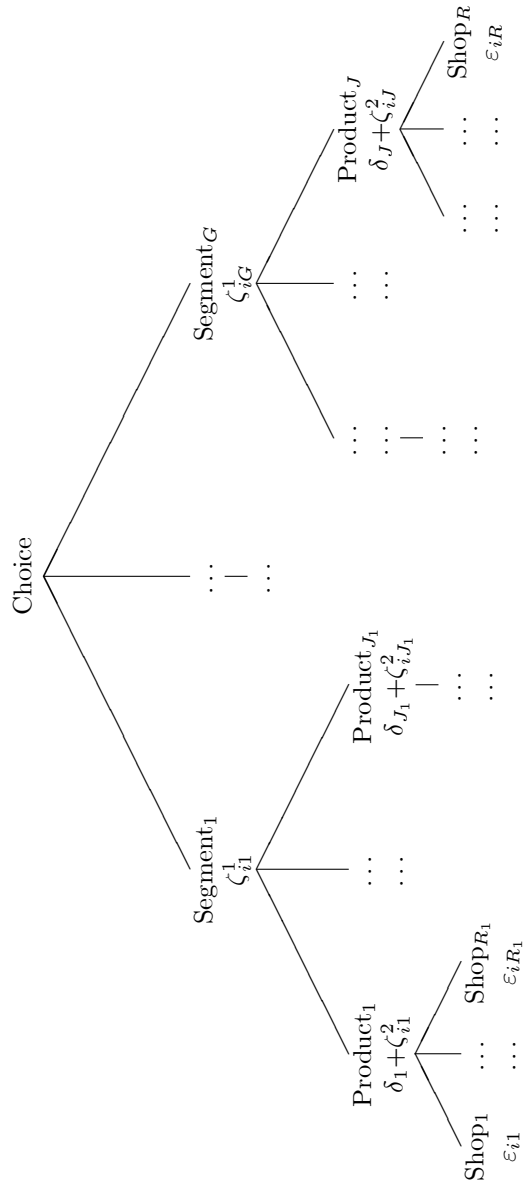
$$u_{ijr} = \delta_j + \zeta_{ij} + \sigma_r \varepsilon_{ijr}$$



$$s_j = \frac{r_j^{\sigma_r} \exp(\delta_j)}{\sum_{l=1}^J r_l^{\sigma_r} \exp(\delta_l)} = \frac{\exp(\delta_j^r)}{\sum_{l=1}^J \exp(\delta_l^r)}$$

Figure 5: Structural Nested Logit Model

$$u_{ijr} = \delta_j + \zeta_{ig}^1 + \sigma_g(\zeta_{ij}^2 + \sigma_r \varepsilon_{ijr})$$



$$s_{j/g} = \frac{\exp\left(\frac{\sigma_r \ln r_j + \delta_j}{\sigma_g}\right)}{\sum_{l \in g} \exp\left(\frac{\sigma_r \ln r_l + \delta_l}{\sigma_g}\right)} \quad s_g = \frac{\left[\sum_{l \in g} \exp\left(\frac{\sigma_r \ln r_l + \delta_l}{\sigma_g}\right)\right]^{\sigma_g}}{\sum_g \left[\sum_{l \in g} \exp\left(\frac{\sigma_r \ln r_l + \delta_l}{\sigma_g}\right)\right]^{\sigma_g}}$$

## 6 Appendix

### 6.A Derivation of equation (5)

For notational convenience in this section we omit subscripts  $i$  and  $t$  and concentrate on the role of products,  $j$ , and stores,  $r$ .

Cardell (1997) shows us that if  $\varepsilon_j$  is an *iid* extreme value random variable with cdf  $F(\varepsilon_j) = e^{-e^{-\varepsilon_j}}$ , and if  $\zeta$  is another random variable with  $\zeta \sim C(\sigma_r)$  and  $0 < \sigma_r < 1$ , then the sum of the two random variables, which we denote with  $\varepsilon_j^1 \equiv (\zeta_j + \sigma_r \varepsilon_j)$ , has itself an extreme value distribution with cdf:

$$F(\varepsilon_j^1) = e^{-\left(e^{-\frac{\varepsilon_j^1}{\sigma_r}}\right)^{\sigma_r}}. \quad (23)$$

Given the *iid* assumption, the cdf in (23) can easily be extended to the following  $J$ -dimensional joint distribution:

$$F(\varepsilon_1^1, \dots, \varepsilon_J^1) = e^{-\left(\sum_{j=1}^J e^{-\frac{\varepsilon_j^1}{\sigma_r}}\right)^{\sigma_r}}.$$

We wish to show the derivation of the logit demand in (5). We begin the derivation by integrating over the distribution of stores carrying product  $j$ . We follow Akiva and Lerhman (1985) and, for each product  $j$ , we compute the expected value of the maximum utility over the stores carrying that product.<sup>21</sup> We find,

$$E[\max(u_{j1}, \dots, u_{jR_j})]. \quad (24)$$

This expectation requires us to recover the probability:

$$\Pr(u_{jr} \geq u_{j1}, \dots, u_{jr} \geq u_{jR_j}),$$

which, consistently with the notation above, can be rewritten as:

$$\Pr(\varepsilon_{j1}^1 \leq \varepsilon_{jr}^1, \dots, \varepsilon_{jR_j}^1 \leq \varepsilon_{jr}^1).$$

Cardell (1997) shows its joint cdf can be formulated as:

$$F(\varepsilon_{j1}^1, \dots, \varepsilon_{jR_j}^1) = e^{-\left(\sum_{s=1}^{R_j} e^{-\varepsilon_{js}^1/\sigma_r}\right)^{\sigma_r}},$$

and benefiting of the *iid* structure we retype it as:

$$F(\varepsilon_{jr}^1) = e^{-R_j^{\sigma_r} \left(e^{-\varepsilon_{jr}^1/\sigma_r}\right)^{\sigma_r}}.$$

---

<sup>21</sup>The relation shops-products can be visualized in the choice tree of Figure 4.

Its pdf is:

$$f(\varepsilon_{jr}^1) = R_j^{\sigma_r} e^{-R_j^{\sigma_r} \left(e^{-\varepsilon_{jr}^1/\sigma_r}\right)^{\sigma_r}} \left(e^{-\varepsilon_{jr}^1/\sigma_r}\right)^{\sigma_r}.$$

We now have all the components required to compute the expected value in (24):

$$E[\max(u_{j1}, \dots, u_{jR_j})] = \int_{-\infty}^{\infty} \varepsilon_{jr}^1 R_j^{\sigma_r} e^{-R_j^{\sigma_r} \left(e^{-\varepsilon_{jr}^1/\sigma_r}\right)^{\sigma_r}} \left(e^{-\varepsilon_{jr}^1/\sigma_r}\right)^{\sigma_r} d\varepsilon_{jr}^1. \quad (25)$$

Using the change of variable  $\vartheta_{jr} = \left(e^{-\varepsilon_{jr}^1/\sigma_r}\right)^{\sigma_r}$  and being aware that  $\frac{d\vartheta_{jr}}{d\varepsilon_{jr}^1} = -\left(e^{-\varepsilon_{jr}^1/\sigma_r}\right)^{\sigma_r}$  and that  $\ln \vartheta_{jr} = -\varepsilon_{jr}^1$ , the integral in (25) simplifies to,

$$\int_{\infty}^0 e^{-R_j^{\sigma_r} \vartheta_{jr}} R_j^{\sigma_r} \ln \vartheta_{jr} d\vartheta_{jr} = - \int_0^{\infty} e^{-R_j^{\sigma_r} \vartheta_{jr}} R_j^{\sigma_r} \ln \vartheta_{jr} d\vartheta_{jr} = \ln R_j^{\sigma_r}.$$

We want to now find the probability that product  $j$  is chosen. The logic is similar to the method just presented above for the store dimension. So we can formulate the joint cdf as,

$$\Pr(\varepsilon_0^1 \leq \delta_j + \ln R_j^{\sigma_r} - \delta_0 - \ln R_0^{\sigma_r} + \varepsilon_j^1, \dots, \varepsilon_J^1 \leq \delta_j + \ln R_j^{\sigma_r} - \delta_J - \ln R_J^{\sigma_r} + \varepsilon_j^1). \quad (26)$$

One can notice (26) accounts for the composite good, 0. Also, we have dropped the subscript  $r$  from the error term, as to give the idea that we have already integrated out the store dimension. We express product  $j$ 's market share as,

$$\begin{aligned} s_j &= \int_{-\infty}^{\infty} \left(e^{-\varepsilon_j^1/\sigma_r}\right)^{\sigma_r} \prod_{l=0}^J e^{-\left(e^{-\frac{\delta_j + \ln R_j^{\sigma_r} - \delta_l - \ln R_l^{\sigma_r} + \varepsilon_j^1}{\sigma_r}}\right)^{\sigma_r}} d\varepsilon_j^1 \\ &= \int_{-\infty}^{\infty} \left(e^{-\varepsilon_j^1/\sigma_r}\right)^{\sigma_r} e^{-\sum_{l=0}^J \left(e^{-\frac{\delta_j + \ln R_j^{\sigma_r} - \delta_l - \ln R_l^{\sigma_r} + \varepsilon_j^1}{\sigma_r}}\right)^{\sigma_r}} d\varepsilon_j^1 \\ &= \int_{-\infty}^{\infty} \left(e^{-\varepsilon_j^1/\sigma_r}\right)^{\sigma_r} e^{-\left(e^{-\varepsilon_j^1/\sigma_r}\right)^{\sigma_r} \sum_{l=0}^J \left(e^{-\frac{\delta_j + \ln R_j^{\sigma_r} - \delta_l - \ln R_l^{\sigma_r}}{\sigma_r}}\right)^{\sigma_r}} d\varepsilon_j^1. \end{aligned} \quad (27)$$

If we define,

$$\lambda_j \equiv \ln \left[ \sum_{l=0}^J \left( e^{-\frac{\delta_j + \ln R_j^{\sigma_r} - \delta_l - \ln R_l^{\sigma_r}}{\sigma_r}} \right)^{\sigma_r} \right],$$



the relation in (27) simplifies to:

$$\begin{aligned} s_j &= \int_{-\infty}^{\infty} \left( e^{-\varepsilon_j^1/\sigma_r} \right)^{\sigma_r} e^{-\left( e^{-\varepsilon_j^1/\sigma_r} \right)^{\sigma_r} e^{\left( \frac{\lambda_j}{\sigma_r} \right)^{\sigma_r}}} d\varepsilon_j^1 \\ &= \int_{-\infty}^{\infty} \left( e^{-\varepsilon_j^1/\sigma_r} \right)^{\sigma_r} e^{-\left( e^{-\left( \varepsilon_j^1 - \lambda_j \right)/\sigma_r} \right)^{\sigma_r}} d\varepsilon_j^1, \end{aligned}$$

and further, if we define  $\eta_j \equiv \varepsilon_j^1 - \lambda_j$ , we are able to derive the market share for product  $j$  as:

$$\begin{aligned} s_j &= \int_{-\infty}^{\infty} \left( e^{-(\eta_j + \lambda_j)/\sigma_r} \right)^{\sigma_r} e^{-(e^{-\eta_j/\sigma_r})^{\sigma_r}} d\eta_j \\ &= e^{-\lambda_j} \int_{-\infty}^{\infty} \left( e^{-\eta_j/\sigma_r} \right)^{\sigma_r} e^{-(e^{-\eta_j/\sigma_r})^{\sigma_r}} d\eta_j, \end{aligned}$$

which has solution

$$s_j = e^{-\lambda_j}. \quad (28)$$

With some trivial algebraic manipulation one can show that equation (28) is the same as our demand equation (5).

## 6.B Some Useful Algebra

### 6.B.1 Logit Specification

We use the product market share derived in equation (5) and compute the following partial derivatives,

$$\begin{aligned} \frac{\partial s_{jt}}{\partial p_{jt}} &= \alpha \left\{ -\frac{\exp(\delta_{jt}^r) \left[ 1 + \sum_{l=1}^{J_t} \exp(\delta_{lt}^r) \right]}{\left[ 1 + \sum_{l=1}^{J_t} \exp(\delta_{lt}^r) \right]^2} + \frac{[\exp(\delta_{jt}^r)]^2}{\left[ 1 + \exp(\sum_{l=1}^{J_t} \delta_{lt}^r) \right]^2} \right\} = -\alpha s_{jt} (1 - s_{jt}) \\ \frac{\partial s_{jt}}{\partial p_{lt}} &= \alpha \frac{\exp(\delta_{jt}^r) \exp(\delta_{lt}^r)}{\left[ 1 + \exp(\sum_{l=1}^{J_t} \delta_{lt}^r) \right]^2} = \alpha s_{lt} s_{jt}, \text{ if } l \neq j, \end{aligned}$$

where  $\delta_{jt}^r \equiv \delta_{jt} + \sigma_r \ln r_{jt}$ , and  $\alpha > 0$ .

The above derivatives can be written in matrix notation as,

$$\frac{\partial}{\partial \mathbf{p}_t} \mathbf{s}_t' = -\alpha \left[ \mathbf{s}_t (\mathbf{i}_t - \mathbf{s}_t) \right] \cdot \mathbf{I}_t + \alpha (\mathbf{s}_t \mathbf{s}_t') \cdot (\mathbf{i}_t \mathbf{i}_t' - \mathbf{I}_t)$$

where, a dot,  $\cdot$ , indicates an operation element by element. The vector of market share  $\mathbf{s}_t$  is  $J_t \times 1$ ,  $\mathbf{p}_t$  is a  $J_t \times 1$  vector of prices,  $\mathbf{i}_t$  is a  $J_t \times 1$  vector of ones, and  $\mathbf{I}_t$  is a  $J_t \times J_t$  identity matrix.

The logit price elasticities are formulated as,

$$\begin{aligned}\frac{\partial s_{jt}}{\partial p_{jt}} \frac{p_{jt}}{s_{jt}} &= -\alpha p_{jt} (1 - s_{jt}) \\ \frac{\partial s_{jt}}{\partial p_{lt}} \frac{p_{lt}}{s_{jt}} &= \alpha p_{lt} s_{lt} \text{ if } l \neq j,\end{aligned}$$

and their matrix notation is,

$$\left[ \frac{\partial}{\partial \mathbf{p}_t} \mathbf{s}_t' \right] \cdot \left[ \mathbf{p}_t (1./\mathbf{s}_t)' \right] = \left\{ -\alpha \left[ \mathbf{p}_t (\mathbf{i}_t' - \mathbf{s}_t') \right] \right\} \cdot \mathbf{I}_t + \alpha \left[ (\mathbf{p}_t \cdot \mathbf{s}_t) \mathbf{i}_t' \right] \cdot (\mathbf{i}_t \mathbf{i}_t' - \mathbf{I}_t).$$

**Logit Counterfactual** Our counterfactual requires that we recompute the underlying price equilibrium in the market. We do not worry here about multiple equilibria. The price equilibrium that we observe is the one we are interested in, and the new equilibrium that we compute is going to be the one of its neighbor.

Our system of demand and pricing equations can be written in compact form as,

$$\mathbf{s}_t^n = \left[ \mathbf{exp}(\boldsymbol{\delta}_t) \mathbf{i}_t' - \mathbf{I}_t \right]^{-1} \mathbf{exp}(\boldsymbol{\delta}_t) \quad (29a)$$

$$\mathbf{p}_t^n = \mathbf{c}_t + \underbrace{\boldsymbol{\Delta}_t^{-1} \mathbf{s}_t^n}_{\mathbf{m}_t^n}, \quad (29b)$$

where the superscript  $n$  indicates the new values, and the mean utilities and marginal costs without superscript  $r$  reference those under full store coverage. The first  $J_t$  equations in (29a) represent our market shares (demand) equations and the subsequent  $J_t$  equations (29b) the pricing equations. Notice that the demand equations are inclusive of the direct changes provided by the counterfactual (i.e.  $r_{jt} = 1, \forall j, t$ ) and the indirect changes provided by the prices. All of them enter via the products mean utility  $\boldsymbol{\delta}_t$ . We then substitute (29a) into (29b) and obtain a  $J_t$  set of nonlinear pricing equations which is the system of equations that has as a solution a new price equilibrium. We use *fsolve* in Matlab to find the solution of this system. The algorithm updates prices (and market shares) until there is convergence to the new equilibrium. Solutions are found for each period  $t$ .

### 6.B.2 Nested Logit Specification

We use the product market shares epitomized in equation (8), and derive the following partial derivatives,

$$\begin{aligned}
\frac{\partial s_{jt}}{\partial p_{jt}} &= \alpha s_{jt} \left[ s_{jt} + \frac{1 - \sigma_g}{\sigma_g} s_{j/g,t} - \frac{1}{\sigma_g} \right], \\
\frac{\partial s_{jt}}{\partial p_{lt}} &= \alpha s_{jt} \left[ s_{lt} + \frac{1 - \sigma_g}{\sigma_g} s_{l/g,t} \right], \text{ if } l \in g \text{ and } l \neq j, \\
\frac{\partial s_{jt}}{\partial p_{lt}} &= \alpha s_{jt} s_{lt}, \text{ if } l \notin g \text{ and } l \neq j,
\end{aligned} \tag{30}$$

which can be abridged in matrix notation as:

$$\begin{aligned}
\left[ \frac{\partial}{\partial \mathbf{p}_t} \mathbf{s}_t' \right] &= \alpha \left[ \mathbf{s}_t \left( \mathbf{s}_t + \frac{1 - \sigma_g}{\sigma_g} \mathbf{s}_{/g,t} - \frac{1}{\sigma_g} \mathbf{i}_t \right)' \right] \cdot \mathbf{I}_t + \\
&+ \alpha \left[ \left( \mathbf{s}_t + \frac{1 - \sigma_g}{\sigma_g} \mathbf{s}_{/g,t} \right) \mathbf{s}_t' \right] \cdot (\mathbf{M}_t - \mathbf{I}_t) + \alpha [(\mathbf{s}_t \mathbf{s}_t')] \cdot (\mathbf{i}_t \mathbf{i}_t' - \mathbf{M}_t).
\end{aligned} \tag{31}$$

Notice that  $\mathbf{M}_t$  is a block diagonal matrix that has along the diagonal sub-matrices of ones for products belonging to the same segment:<sup>22</sup>

$$\mathbf{M}_t \equiv \begin{bmatrix} \mathbf{M}_{1t} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_{Gt} \end{bmatrix}.$$

The resulting price elasticities are,

$$\begin{aligned}
\frac{\partial s_{jt}}{\partial p_{jt}} \frac{p_{jt}}{s_{jt}} &= \alpha p_{jt} \left[ s_{jt} + \frac{1 - \sigma_g}{\sigma_g} s_{j/g,t} - \frac{1}{\sigma_g} \right] \\
\frac{\partial s_{jt}}{\partial p_{lt}} \frac{p_{lt}}{s_{jt}} &= \alpha p_{lt} \left[ s_{lt} + \frac{1 - \sigma_g}{\sigma_g} s_{l/g,t} \right] \text{ if } l \in g \text{ and } l \neq j, \\
\frac{\partial s_{jt}}{\partial p_{lt}} \frac{p_{lt}}{s_{jt}} &= \alpha p_{lt} s_{lt} \text{ if } l \notin g \text{ and } l \neq j,
\end{aligned}$$

and their matrix notation is:

$$\begin{aligned}
\left[ \frac{\partial}{\partial \mathbf{p}_t} \mathbf{s}_t' \right] \cdot [\mathbf{p}_t (1./\mathbf{s}_t)] &= \alpha \left[ \mathbf{p}_t \left( \mathbf{s}_t + \frac{1 - \sigma_g}{\sigma_g} \mathbf{s}_{/g,t} - \frac{1}{\sigma_g} \mathbf{i}_t \right)' \right] \cdot \mathbf{I}_t + \\
&+ \alpha \left\{ \left[ \mathbf{p}_t \cdot \left( \mathbf{s}_t + \frac{1 - \sigma_g}{\sigma_g} \mathbf{s}_{/g,t} \right) \right] \mathbf{i}' \right\} \cdot (\mathbf{M}_t - \mathbf{I}_t) + \\
&+ \alpha [(\mathbf{p}_t \cdot \mathbf{s}_t) \mathbf{i}'] \cdot (\mathbf{i}_t \mathbf{i}_t' - \mathbf{M}_t).
\end{aligned} \tag{32}$$

---

<sup>22</sup>The matrix  $\mathbf{M}_t$  requires the products to be ordered by segments. Notice that matrices  $\mathbf{M}_{gt}$  along the diagonal have  $\dim(\mathcal{J}_{gt})$ , where  $\dim(\mathcal{J}_{gt})$  is the number of products entering segment  $g$  in period  $t$ .

**Nested Logit Counterfactual** Given that a nested logit can be expressed as the product of two logits,  $s_{jt} = s_{j/g,t} s_{gt}$ , the technique used to compute our counterfactual is similar to that outlined in subsection (6.B.1) with the addition that, this time, we also need to update the within segment market shares and be aware of the nested logit structure.

### 6.B.3 Random Coefficients Nested Logit Specification

We average the product market shares in (13), and derive the following partial derivatives,

$$\begin{aligned}\frac{\partial s_{jt}}{\partial p_{jt}} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ijt}}{\partial p_{jt}} = \frac{1}{ns} \sum_{i=1}^{ns} \alpha_i s_{ijt} \left[ s_{ijt} + \frac{1 - \sigma_g}{\sigma_g} s_{i,j/g,t} - \frac{1}{\sigma_g} \right], \quad (33) \\ \frac{\partial s_{jt}}{\partial p_{lt}} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ijt}}{\partial p_{lt}} = \alpha_i s_{ijt} \left[ s_{ilt} + \frac{1 - \sigma_g}{\sigma_g} s_{i,l/g,t} \right], \text{ if } l \in g \text{ and } l \neq j, \\ \frac{\partial s_{jt}}{\partial p_{lt}} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ijt}}{\partial p_{lt}} = \alpha_i s_{ijt} s_{ilt}, \text{ if } l \notin g \text{ and } l \neq j,\end{aligned}$$

and its compact form is an immediate adaptation of (31), such as:

$$\begin{aligned}\left[ \frac{\partial}{\partial \mathbf{p}_t} \mathbf{s}_t' \right] &= \frac{1}{ns} \sum_{i=1}^{ns} \alpha_i \left\{ \left[ \mathbf{s}_{it} \left( \mathbf{s}_{it} + \frac{1 - \sigma_g}{\sigma_g} \mathbf{s}_{i/g,t} - \frac{1}{\sigma_g} \mathbf{i}_t \right) \right]' \cdot \mathbf{I}_t + \right. \\ &\quad \left. + \left[ \left( \mathbf{s}_{it} + \frac{1 - \sigma_g}{\sigma_g} \mathbf{s}_{i/g,t} \right) \mathbf{s}_t' \right] \cdot (\mathbf{M}_t - \mathbf{I}_t) + [(\mathbf{s}_{it} \mathbf{s}_{it}')] \cdot (\mathbf{i}_t \mathbf{i}_t' - \mathbf{M}_t) \right\}.\end{aligned}$$

The formulas for the price elasticities follow trivially, and for that reason we decided to skip their representation.

**Random Coefficients Nested Logit Counterfactual** The way we compute our counterfactual for the random coefficient (logit) model is alike those presented in the previous subsections (6.B.1 and 6.B.2), with the additional difficulty that we now have to deal with simulated individual probabilities. Hence the markup function relies on the partial derivatives displayed in equation (33). The underlying logic is however unchanged: *Fsolve* in Matlab will deliver the new equilibrium solution to us.

## 6.C Endogeneity: Simultaneity and Omitted Variable

The identification of a Least Squares estimator in the case of stochastic explanatory variables typically avails of the conditional independence assumption, that

is, the error term has to be mean zero once conditioned on a set of exogenous variables. This was condition (22) in our paper.

We now write the following relation:

$$y = X\beta + \varepsilon,$$

and put forward the following conditional independence assumption  $E(\varepsilon|X) = 0$ , but assume that only a subset  $X_1$  of the variables in  $X \equiv [X_1, X_2]$  is observed. The remaining variables  $X_2$ , are omitted. The unobservability of  $X_2$  generates a violation of the conditional independence assumption  $E(\varepsilon|X) = 0$ , and thus Least Squares estimated parameters are biased. The issue is that the true model is,

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon,$$

but we estimate by Least Squares the model,

$$y = X_1\beta_1 + \tilde{\varepsilon},$$

where

$$\tilde{\varepsilon} = X_2\beta_2 + \varepsilon.$$

The Least Squares estimator of the model would lead to,

$$\hat{\beta}_1 = (X_1'X_1)^{-1} (X_1'y).$$

Now, if the true model is correctly specified, condition  $E(\varepsilon|X) = 0$  implies  $E(\varepsilon|X_1) = 0$ , so that,

$$E(\hat{\beta}_1|X) = \beta_1 + \underbrace{E\left[(X_1'X_1)^{-1} (X_1'X_2) \beta_2|X\right]}_{\text{Omitted Variables Bias}}.$$

In this case the bias is generated entirely by the omitted (unobserved) variables  $X_2$ . On the other side, if the model suffers of a combination of simultaneity and omitted variables, the condition  $E(\varepsilon|X_1) = 0$  is itself violated so that the total bias would be,

$$E(\hat{\beta}_1|X) = \beta_1 + \underbrace{E\left[(X_1'X_1)^{-1} (X_1'X_2) \beta_2|X\right]}_{\text{Omitted Variables Bias}} + \underbrace{E\left[(X_1'X_1)^{-1} (X_1'\varepsilon)|X\right]}_{\text{Simultaneity Bias}}.$$

Of course summing two sources of bias can generate the following scenarios: A larger (in absolute value) total bias, if both sources of biases co-move, i.e., have the same sign, or a smaller (in absolute value) total bias, if the two sources of bias move in different directions (have different signs).

We now define  $Z \equiv [Z_1, X_2]$  and assume that  $Z_1$  is a set of instruments that solves the simultaneity bias, so that  $E(\varepsilon|Z_1) = 0$ . This confines the bias to be only produced by omitted variables, and to depend on the relation between the set of instruments  $Z_1$  and the omitted variables  $X_2$ :

$$E(\hat{\beta}_1|X) = \beta_1 + \underbrace{E\left[\left(Z_1'X_1\right)^{-1}\left(Z_1'X_2\right)\beta_2|Z\right]}_{\text{New Omitted Variables Bias}}.$$

An important by-product of this instrumentation is that, should  $Z_1$  and  $X_2$  be uncorrelated, the omitted variables bias would vanish. In other words, it would be possible to search for ideal instruments that would correct simultaneously both the simultaneity and omitted variables bias, alternatively one should either find a good proxy of  $X_2$ , or look for new data that have  $X_2$ .

## 6.D Gradient of the Objective Function Associated to Non-linear Parameters in the Random Coefficient Nested Logit Model

If we average the individual probabilities displayed in equation (12) we obtain the market share of product  $j$  in period  $t$ ,

$$s_{jt} = \frac{1}{ns} \sum_{i=1}^{ns} \frac{D_{igt}^{\sigma_g}}{\sum_{s=0}^G D_{ist}^{\sigma_g}} \frac{\exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right)}{D_{igt}},$$

whose components have been defined in equations (13) and (14). We follow the appendix of Nevo (2000a) and derive the various components that enter the demand side gradient:

1a) The partial derivative of the market share with respect to own mean utility,

$$\begin{aligned}
\frac{\partial s_{jt}}{\partial \delta_{jt}} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ijt}}{\partial \delta_{jt}} \\
&= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\sigma_g D_{igt}^{\sigma_g-1} \exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right) (\sigma_g)^{-1} \sum_{s=0}^G D_{ist}^{\sigma_g} - D_{igt}^{\sigma_g} \sigma_g D_{igt}^{\sigma_g-1} \exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right) (\sigma_g)^{-1}}{\left(\sum_{s=0}^G D_{ist}^{\sigma_g}\right)^2} s_{i,j/g,t} + \\
&\quad + \frac{1}{ns} \sum_{i=1}^{ns} s_{igt} \frac{\exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right) (\sigma_g)^{-1} D_{igt} - \exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right) \exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right) (\sigma_g)^{-1}}{D_{igt}^2} \\
&= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right)}{D_{igt}} \frac{D_{igt}^{\sigma_g} \sum_{s=0}^G D_{ist}^{\sigma_g} - D_{igt}^{\sigma_g} D_{igt}^{\sigma_g}}{\left(\sum_{s=0}^G D_{ist}^{\sigma_g}\right)^2} s_{i,j/g,t} + \\
&\quad + \frac{1}{ns} \sum_{i=1}^{ns} \frac{s_{igt}}{\sigma_g} \frac{\exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right) D_{igt} - \exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right) \exp\left(\frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g}\right)}{D_{igt}^2} \\
&= \frac{1}{ns} \sum_{i=1}^{ns} s_{i,j/g,t} s_{igt} (1 - s_{igt}) s_{i,j/g,t} + \frac{1}{ns} \sum_{i=1}^{ns} \frac{1}{\sigma_g} s_{igt} s_{i,j/g,t} (1 - s_{i,j/g,t}) \\
&= \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} (1 - s_{igt}) s_{i,j/g,t} + \frac{1}{ns} \sum_{i=1}^{ns} \frac{1}{\sigma_g} s_{ijt} (1 - s_{i,j/g,t}) \\
&= \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} \left[ (1 - s_{igt}) s_{i,j/g,t} + \frac{(1 - s_{i,j/g,t})}{\sigma_g} \right] \\
&= \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} \left[ (1 - s_{igt}) s_{i,j/g,t} + \frac{(1 - s_{i,j/g,t})}{\sigma_g} \right] \\
&= \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} \left[ -s_{ijt} - s_{i,j/g,t} \frac{1 - \sigma_g}{\sigma_g} + \frac{1}{\sigma_g} \right].
\end{aligned}$$

We recognize that the above partial derivative is the own price partial derivative displayed in equation (30), but without the additional term  $(\partial \delta_{jt} / \partial p_{jt}) = -\alpha$ . Hence we can fit the partial derivatives in (30) to our case as follows:

1b) The partial derivative of market share with respect to the mean utilities of products in the same segment is:

$$\frac{\partial s_{jt}}{\partial \delta_{mt}} = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ijt}}{\partial \delta_{mt}} = \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} \left( -s_{ijt} - s_{i,j/g,t} \frac{1 - \sigma_g}{\sigma_g} \right),$$

1c) The partial derivative of market share with respect to the mean utilities

of products in other segments is:

$$\frac{\partial s_{jt}}{\partial \delta_{mt}} = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ijt}}{\partial \delta_{mt}} = \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} (-s_{ijt}).$$

We now add the partial derivatives of market shares with respect to the nonlinear parameters, as follows:

2) The partial derivative of market share with respect to  $\sigma_g$  is given by:

$$\begin{aligned} \frac{\partial s_{jt}}{\partial \sigma_g} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ijt}}{\partial \sigma_g} \\ &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\left( D_{igt}^{\sigma_g} (\ln D_{igt}) \left[ \sum_{l \in J_g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) (\sigma_g)^{-2} \right] \sum_{s=0}^G D_{ist}^{\sigma_g} - \right.}{\left( \sum_{s=0}^G D_{ist}^{\sigma_g} \right)^2} \left. - D_{igt}^{\sigma_g} \sum_{s=1}^G \left\{ D_{ist}^{\sigma_g} (\ln D_{ist}) \left[ \sum_{l \in J_s} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) (\sigma_g)^{-2} \right] \right\} \right) s_{i,j/g,t} + \\ &\quad + \frac{1}{ns} \sum_{i=1}^{ns} s_{igt} \frac{\left\{ \exp \left( \frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g} \right) (\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p) (\sigma_g)^{-2} D_{igt} - \right.}{D_{igt}^2} \left. - \exp \left( \frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g} \right) \left[ \sum_{l \in J_g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) (\sigma_g)^{-2} \right] \right\}}{D_{igt}^2} \\ &= \frac{1}{ns} \sum_{i=1}^{ns} \left\{ \frac{s_{igt} (\ln D_{igt}) \left[ \sum_{l \in J_g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) (\sigma_g)^{-2} \nu_i^g \right] -}{-s_{igt} \sum_{s=1}^G \left\{ s_{ist} (\ln D_{ist}) \left[ \sum_{l \in J_s} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) (\sigma_g)^{-2} \nu_i^g \right] \right\}} \right\} s_{i,j/g,t} + \\ &\quad + \frac{1}{ns} \sum_{i=1}^{ns} s_{igt} \left[ s_{i,j/g,t} (\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p) (\sigma_g)^{-2} - s_{i,j/g,t} \left( \sum_{l \in J_g} s_{i,l/g,t} (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) \right) (\sigma_g)^{-2} \right] \\ &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{s_{ijt}}{(\sigma_g)^2} \left\{ \frac{(\ln D_{igt}) \left[ \sum_{l \in J_g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) \right] -}{- \sum_{s=1}^G \left\{ s_{ist} (\ln D_{ist}) \left[ \sum_{l \in J_s} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) \right] \right\}} \right\} + \\ &\quad + \frac{1}{ns} \sum_{i=1}^{ns} \frac{s_{ijt}}{(\sigma_g)^2} \left[ \sigma_r \ln r_{jt} + \delta_{jt} - \sum_{l \in J_g} s_{i,l/g,t} (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) \right] \\ &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{s_{ijt}}{(\sigma_g)^2} \left\{ \frac{(\ln D_{igt}) \left[ \sum_{l \in J_g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) \right] -}{- \sum_{s=1}^G \left\{ s_{ist} (\ln D_{ist}) \left[ \sum_{l \in J_s} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) \right] \right\}} \right\} + \\ &\quad + \left[ \delta_{jt}^r - \sigma_p p_{jt} \nu_i^p - \sum_{l \in J_g} s_{i,l/g,t} (\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p) \right] \end{aligned}$$

3) Similarly, the partial derivatives of market share with respect to the two price parameters have to be taken separately:



i) With respect to the standard deviation parameter,

$$\begin{aligned}
\frac{\partial s_{jt}}{\partial \sigma_p} &= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\partial s_{ijt}}{\partial \alpha} \\
&= \frac{1}{ns} \sum_{i=1}^{ns} \frac{\left( \begin{array}{c} \sigma_g D_{igt}^{\sigma_g-1} \left[ \sum_{l \in J_g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (-p_{lt} \nu_i^p) (\sigma_g)^{-1} \right] \sum_{s=0}^G D_{ist}^{\sigma_g} - \\ - D_{igt}^{\sigma_g} \sum_{s=1}^G \left\{ \sigma_g D_{ist}^{\sigma_g-1} \left[ \sum_{l \in J_s} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (-p_{lt} \nu_i^p) (\sigma_g)^{-1} \right] \right\} \end{array} \right)}{\left( \sum_{s=0}^G D_{ist}^{\sigma_g} \right)^2} s_{i,j/g,t} + \\
&\quad + \frac{1}{ns} \sum_{i=1}^{ns} s_{igt} \frac{\left\{ \begin{array}{c} \exp \left( \frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g} \right) (-p_{jt} \nu_i^p) (\sigma_g)^{-1} D_{igt} - \\ - \exp \left( \frac{\delta_{jt}^r - \sigma_p p_{jt} \nu_i^p}{\sigma_g} \right) \left[ \sum_{l \in J_g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) (-p_{lt} \nu_i^p) (\sigma_g)^{-1} \right] \end{array} \right\}}{D_{igt}^2} \\
&= \frac{1}{ns} \sum_{i=1}^{ns} \left\{ \begin{array}{c} -s_{igt} \sigma_g D_{igt}^{-1} \left[ \sum_{l \in J_g} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) p_{lt} (\sigma_g)^{-1} \right] + \\ + s_{igt} \sum_{s=1}^G \left\{ s_{ist} \sigma_g D_{ist}^{-1} \left[ \sum_{l \in J_s} \exp \left( \frac{\delta_{lt}^r - \sigma_p p_{lt} \nu_i^p}{\sigma_g} \right) p_{lt} (\sigma_g)^{-1} \right] \right\} \end{array} \right\} \nu_i^p s_{i,j/g,t} + \\
&\quad + \frac{1}{ns} \sum_{i=1}^{ns} \nu_i^p s_{igt} \left[ -s_{i,j/g,t} p_{jt} (\sigma_g)^{-1} + s_{i,j/g,t} \left( \sum_{l \in J_g} s_{i,l/g,t} p_{lt} \right) (\sigma_g)^{-1} \right] \\
&= \frac{1}{ns} \sum_{i=1}^{ns} \nu_i^p s_{ijt} \left\{ \begin{array}{c} - \left[ \sum_{l \in J_g} s_{i,l/g,t} p_{lt} \right] + \\ + \sum_{s=1}^G s_{ist} \left[ \sum_{l \in J_s} s_{i,l/g,t} p_{lt} \right] \end{array} \right\} + \\
&\quad + \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} \frac{\nu_i^p}{(\sigma_g)} \left[ -p_{jt} + \left( \sum_{l \in J_g} s_{i,l/g,t} p_{lt} \right) \right] \\
&= \frac{1}{ns} \sum_{i=1}^{ns} \nu_i^p s_{ijt} \left\{ \begin{array}{c} - \left[ \sum_{l \in J_g} s_{i,l/g,t} p_{lt} \right] + \\ + \sum_{s=1}^G \left\{ s_{ist} \left[ \sum_{l \in J_s} s_{i,l/g,t} p_{lt} \right] \right\} - \\ - (\sigma_g)^{-1} \left[ p_{jt} - \sum_{l \in J_g} s_{i,l/g,t} p_{lt} \right] \end{array} \right\},
\end{aligned}$$

ii) With respect to the mean parameter,

$$\frac{\partial s_{jt}}{\partial \alpha} = \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} \left\{ \begin{array}{c} - \left[ \sum_{l \in J_g} s_{i,l/g,t} p_{lt} \right] + \\ + \sum_{s=1}^G \left\{ s_{ist} \left[ \sum_{l \in J_s} s_{i,l/g,t} p_{lt} \right] \right\} - \\ - (\sigma_g)^{-1} \left[ p_{jt} - \sum_{l \in J_g} s_{i,l/g,t} p_{lt} \right] \end{array} \right\}.$$

Those above partial derivatives are the main components that one needs to derive the demand side gradient analytically.

## 6.E Instruments Selection

Before devoting attention to our choice of instruments, we briefly introduce our four step procedure. In step one we regress our endogenous variables, over

different sets of instruments, to see which instruments have more power. In step two, we select the significant instruments, compute a demand side GMM and verify whether the instruments have some validity. Step three avails of the estimated demand primitives from the previous step, deduces estimated marginal costs, and assesses the power on supply. The final step checks for the validity of the instruments for those endogenous variables that directly, or via functional forms, enter the pricing equation GMM. The idea of this process is to explore the power and validity of the instruments separately for demand and supply, as it is likely to be beneficial for the choice of instruments in the joint estimate of demand and supply.

We now detail the process of choosing the instruments. We regress the endogenous variables listed in the first row of Table 11, individually on each set of instruments. Instrumentation includes the BLP ( $z_1 - z_6$ ) and the Hausman and Taylor ( $z_7 - z_8$ ) instruments. We begin by discussing the power of instruments on endogenous demand side variables. Column two of Table 11 exhibits possible instruments for the price variable. The set of instruments displays decent explanatory power. The  $F$ -statistic of 5.65 is significant. Based on this information we suggest  $z_1$ ,  $z_2$ ,  $z_5$ , and  $z_7$  as potential instruments for the endogeneity of prices. Next, because price is the only endogenous variable in the logit specification, we can immediately use those instruments to estimate a GMM logit demand. The results are reported in Table 12. The first column shows a non significant value of the price parameter, and a high value for the Hansen  $J$  overestimation test (25.18). Seemingly our selected instruments are not too helpful in identifying the price parameter. Furthermore, the estimation of the price parameter is not too far off the OLS biased alternative, and thus it generates a similar impressive number of deduced negative marginal costs. With this in mind, we conclude that we have inadequate instruments to estimate a single demand side in a logit specification when store coverage is unobserved. However, we warn that this conclusion only applies to the single equation demand estimation.

We proceed to add the role of store coverage in the estimation of the logit specification. First we attempt to validate the use of the  $AR$  proxy. Looking at column five of Table 11, we observe that the introduction of the proxy does not change our choice of best instruments, but presents a slightly lower  $F$ -test. The GMM demand estimation does provide evidence of a slighter minor improvement in the overidentification test, but overall the estimated coefficients and explanatory power are non satisfactory. We move on to use actual store coverage. Since *a priori* one cannot exclude this variable from being endogenous,

we need to separate the case where store coverage is endogenous from the one that it is exogenous. We begin the process by assuming the variable to be endogenous, and then test for its endogeneity. Additionally, we must investigate potential instruments for store coverage, and the options are illustrated in the final column of Table 11. Clearly  $z_1$ , will be a crucial instrument to capture the endogeneity of store coverage, as it is significant for store coverage and not for price (the other endogenous variable). Column three of the GMM demand estimation in Table 12 depicts the estimates when store coverage is treated as endogenous. The p-value of the Hansen C test at the bottom of the table suggests that the variable should be treated as exogenous. Hence, we re-estimate the GMM demand equation, this time assuming store coverage as exogenous and write down the results in column four. From the table it emerges that the demand parameter is identified. The number of deduced negative marginal costs is now declining, however we still have that the parameter associated to store coverage is overestimated (exceeds its theoretical bound of one); probably due to the omittance of a full nested logit structure.

We now turn to the nested logit specification and start the previous procedure again with store coverage unobserved. This time we need to add the within segment market share endogenous variable. Column three in Table 11 unveils  $z_1$ ,  $z_2$ ,  $z_4$ , and  $z_7$  to be powerful in explaining this endogeneity. We rely on  $z_4$  as good instrument to identify the within segment coefficient, as it is shown to be weakly significant in explaining the within segment market share regression, and not significant in the price estimation. The estimation of the GMM demand, for which results are detailed in Table 12, indicates significant overestimations of the price and within segment market share parameters. This suggests a potential misspecification, probably due to the unobservability of store coverage.

We now convey our attention to the nested logit specification under the *AR* proxy scenario. From Table 11 we witness the strong power of  $z_1$ ,  $z_2$ , and  $z_7$  in explaining the log of within segment market share, so combined with the aforementioned chosen instruments for the price variable, we end up using  $z_1$ ,  $z_2$ ,  $z_4$ ,  $z_5$ , and  $z_7$  for the demand side estimation. Its results are documented in Table 12. We see that the introduction of the *AR* proxy corrects for the overestimation of the price (in absolute value) and for within segment market share parameters, but the low p-value reported in the overidentification test still signals a certain weakness of the instrumentation.

Our final step is to add the observability of store coverage and verify how that contributes to our demand nested logit estimation. We select the good

instruments from the regressions printed in the last column of Table 11 and embed them in the GMM demand estimation. Next, we estimate demand assuming store coverage as endogenous. Then, we re-estimate it by imposing store coverage as exogenous. Based on the value of a Hansen  $C$  test we decide whether or not store coverage is endogenous in this second specification. The Hansen  $C$  test p-value expressed in column seven of Table 12 tells us that store coverage should be considered exogenous. Summarizing the other main results in the table: The GMM demand estimation that uses store coverage produces better results than the one that avails of the AR proxy. However, the instrumentation still presents a low validity of the instruments (as indicated by the Hansen  $J$  test); nevertheless comparing it with the previous specification, we see movements in the right direction.

As discussed at the beginning of this section we use the estimated demand primitives to recover an estimate of the marginal costs. So, we can shift to stages three and four of our procedure. In stage three we regress the following endogenous variables: marginal cost for the logit specification  $c$ , marginal cost for the nested logit specification  $nc$ , and the demand side variables market share  $s$ , and within segment market share  $s_{j/g}$ . This step of the procedure helps us in choosing the powerful instruments that are to be rooted in the GMM “supply” estimation.

We begin this “supply” side digression by analyzing the estimates of the logit specification with unobserved store coverage. Column two of Table 13 shows that there is only one instrument good at explaining the market share  $s$ , and that is  $z_1$ . Based on this information we choose  $z_1$  as an instrument for the cost equation, so that the cost side in this case is just identified. Its  $IV$  estimation is outlined in Table 14 and the results are not too encouraging. This is most probably a follow-up of the bias in the price parameter of the GMM demand estimation, as it enters the market power function, and has therefore a direct impact in our derivation of marginal costs. Next, also for the cost side, we shed light on the role of the AR proxy. The instrumentation is the same as we have outlined for logit without store coverage (only one good instrument, which is  $z_1$ ). As for the GMM “supply” estimation, column two of Table 14 shows the estimates have not improved from those reported in column one (again, possibly due to the bias in the demand GMM estimation). The results get better if we observe store coverage. In this case our pick of instruments leans towards  $z_3$ ,  $z_4$ , and  $z_6$ . The GMM “supply” estimation is now satisfactory. The instruments pass the validity test and show interesting results. Covering more stores affects the marginal cost of production in a significant way. In addition, the longer the

period to stock out, the lower the marginal cost of production, therefore this is a good proxy for transportation cost.

We continue our analysis of stage 3 and 4 using the nested logit. The regression of the within segment market share is presented in column four of Table 13. The significant instruments are  $z_1$ ,  $z_4$ , and  $z_7$  and they will all be part of the GMM “supply” estimation.

Following the method previously set out, we start with the case where store coverage is unobservable. The GMM estimate shows a good quality of instruments, as indicated by the Hansen  $J$  test. The variable months to stock out is not significant here, but it carries the right sign. The next step is the familiar attempt to utilize the  $AR$  proxy. In this scenario we pick only two powerful instruments:  $z_4$ , and  $z_7$ . We are again just identified. The parameters of the GMM estimation are emboldening. They are of the right sign and significant. Again, finally we account for the true store coverage. In this event the choice of the instruments points to  $z_3$ ,  $z_4$ ,  $z_6$ . The results of the GMM are encouraging, both in the sign and significance of the relevant parameters in terms of months to stock out and store coverage.



Table 12: GMM demand estimation: Dependent Variable Log Normalized Market Shares

VARIABLES	AR			SC end	SC ex	N	N AR	N SC end	NSC ex
	Logit						Nested Logit		
$\beta_1$ [cons]	-6.064*** (2.016)	-3.702* (2.154)	-7.415*** (1.004)	-2.328** (1.047)		7.497** (3.449)	5.726 (3.672)	9.030*** (3.384)	2.513* (1.284)
$\beta_2$ [inv]	-20.994*** (3.638)	-22.932*** (4.015)	-8.759*** (2.457)	-7.202*** (1.963)		-2.801 (4.802)	-4.660 (5.264)	2.543 (2.874)	-3.628*** (1.127)
$-\alpha$ [p]	0.554 (1.665)	1.281 (1.785)	-0.945 (0.814)	3.337*** (0.864)		8.284*** (2.135)	6.848*** (2.374)	9.035*** (2.182)	5.339*** (0.905)
$\sigma^r$ [ $r^\dagger$ ]		1.938*** (0.535)	1.082*** (0.056)	1.075*** (0.019)			0.556*** (0.550)	0.209 (0.183)	0.437*** (0.102)
$(1 - \sigma^g)$ [ $is_g$ ]						1.018*** (0.250)	0.910*** (0.280)	1.147*** (0.229)	0.641*** (0.099)
<b>STATISTICS</b>									
N.	3644	3644	3644	3644		3644	3644	3644	3644
N. mc<0	3590	1800	3644	53		10	0	0	0
$R^2$	0.359	0.361	0.809	0.783		0.365	0.541	0.197	0.725
Hansen C p-val lrj	.	.	0.639	.		.	.	0.001	.
Hansen J stat	25.180	7.952	85.870	41.120		8.133	12.330	8.137	12.440
Hansen J p-val	1.42e-05	0.047	0	8.87e-08		0.043	0.006	0.043	0.029

Robust Standard errors in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1, † AR proxy in the AR specifications.  
Not reported firm, segment, and (package x time) fixed effects.

Table 13: Power of the supply IV

VARIABLES	c	nc	s	s/g	c AR	nc AR	s AR	s/g AR	c SC	nc SC	s SC	s/g SC
cons	-0.757*** (0.077)	1.187*** (0.412)	0.011*** (0.003)	0.248** (0.111)	0.306*** (0.079)	1.078*** (0.079)	0.011*** (0.003)	0.253** (0.120)	0.330** (0.137)	0.567*** (0.141)	0.017*** (0.005)	0.462** (0.191)
inv	0.008 (0.116)	-0.422 (0.619)	-0.032*** (0.004)	-1.864*** (0.167)	-0.030 (0.110)	-0.003 (0.111)	-0.032*** (0.004)	-1.863*** (0.168)	-0.097 (0.113)	-0.068 (0.116)	-0.010** (0.004)	-0.555*** (0.157)
$I_t^\dagger$					-0.189 (0.499)	-0.261*** (0.504)	-0.005 (0.019)	-0.119 (0.102)	-0.002 (0.002)	-0.006** (0.002)	0.002*** (0.000)	0.090*** (0.003)
$z_1$	-0.028*** (0.002)	-0.002 (0.008)	0.000** (0.000)	0.005** (0.002)	-0.015*** (0.003)	-0.007** (0.003)	0.000 (0.000)	0.005 (0.004)	-0.003 (0.002)	-0.002 (0.002)	0.000 (0.000)	0.002 (0.003)
$z_2$	0.000*** (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)
$z_3$												
$z_4$	-0.001** (0.001)	0.001 (0.003)	-0.000 (0.000)	-0.002* (0.001)	-0.002 (0.003)	-0.002 (0.003)	-0.000 (0.000)	-0.001 (0.004)	0.005*** (0.002)	0.005*** (0.002)	-0.000* (0.000)	-0.004** (0.002)
$z_5$	0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000** (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000*** (0.000)	-0.000 (0.000)	-0.000 (0.000)
$z_6$												
$z_7$	-0.041*** (0.010)	0.084 (0.054)	0.000 (0.000)	-0.084*** (0.014)	-0.041*** (0.009)	-0.029*** (0.010)	0.000 (0.000)	-0.084*** (0.014)	-0.041*** (0.009)	-0.037*** (0.010)	0.000 (0.000)	-0.081*** (0.013)
<b>STATISTICS</b>												
N.	3644	3644	3644	3644	3644	3644	3644	3644	3644	3644	3644	3644
$R^2$	0.852	0.193	0.286	0.280	0.830	0.818	0.286	0.280	0.826	0.819	0.348	0.420
F test instr	61.77	0.784	11.57	29.09	17.72	3.887	9.925	24.93	8.943	7.128	45.24	117.5

Standard errors in parentheses; \*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1, † AR proxy in the AR specifications.

Not reported firm, segment, and (package x time) fixed effects.



Table 14: GMM Supply Estimation

VARIABLES	c	c AR	c SC	nc	nc AR	nc SC
$\gamma_1$ [cons]	0.854 (0.708)	1.098*** (0.178)	1.127*** (0.065)	1.138*** (0.125)	1.402*** (0.108)	1.393*** (0.076)
$\gamma_2$ [inv]	-8.828* (4.746)	-2.244*** (0.824)	-0.511** (0.215)	-1.165 (0.725)	-0.523 (0.386)	-0.414* (0.216)
$\gamma_3$ [ $\pi^{\dagger}$ ]		0.411*** (0.123)	0.055*** (0.018)		0.056 (0.080)	0.041** (0.019)
s (end)	-274.295** (131.740)	-68.174*** (20.804)	-37.697*** (10.567)	8.839 (21.102)	-29.848*** (11.341)	-45.891*** (12.748)
$s/g$ (end)				-0.672* (0.385)	0.244* (0.131)	0.258** (0.118)
<b>STATISTICS</b>						
N.	3644	3644	3644	3644	3644	3644
$R^2$	.	.	0.497	0.178	0.640	0.406
Hansen J p-val	.	.	0.920	0.260	.	0.956
Hansen J stat	.	.	0.167	1.269	.	0.0893

Robust Standard errors in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

$\dagger$  AR proxy in the AR specifications. Not reported firm, segment, and (package x time) fixed effects.

A dot in the overidentification J test is due to the lack of overidentifying restrictions.