

# Online Appendix

## Multi-Category Competition and Market Power: A Model of Supermarket Pricing

Øyvind Thomassen, Howard Smith,  
Stephan Seiler and Pasquale Schiraldi

### APPENDIX A: CATEGORY DEFINITIONS

TNS assigns to each transaction the variable “Retailer Share Track (RST) Market Code” that correspond to 269 narrowly defined product groups. We define our eight categories as follows where the names of product groups (including abbreviations) are those of TNS.

- 1) *Bakery*: Ambient Pizza Bases, Ambient Cakes and Pastries, Ambient Christmas Pudding, Ambient Sponge Puddings, Canned Rice Puddings, Childrens Biscuits, Chilled Breads, Chilled Cakes, Chilled Desserts, Chilled Pizza and Bases, Crackers & Crispbreads, Everyday Biscuits, Fresh/Chilled Pastry, Frozen Bread, Frozen Savoury Bakery, Healthier Biscuits, Morning Goods, Savoury Biscuits, Seasonal Biscuits, Tinned Sponge Puddings, Toaster Pastries, Total Bread.
- 2) *Dairy*: Butter, Defined Milk and Cream Products, Fresh Cream, Fromage Frais, Instant Milk, Margarine, Total Cheese, Total Ice Cream, Yoghurt, Yoghurt Drinks And Juices.
- 3) *Drink*: Ambient One Shot Drinks, Ambient Fruit or Yoghurt Juice and Drnk, Beer and Lager, Bottled Colas, Bottled Lemonade, Bottled Other Flavours, Bottled Shandies, Canned Colas, Canned Lemonade, Canned Other Flavours, Canned Shandies, Chilled One Shot Drinks, Cider, Fabs, Food Drinks, Fortified Wines, Ginger Ale, Lemon and Lime Juices, Mineral Water, Soda Water, Sparkling Wine, Spirits, Tonic Water, Wine.
- 4) *Dry Grocery* : Ambient Condiments, Ambient Slimming Products, Ambient Vegetarian Products, Artificial Sweeteners, Breakfast Cereals, Chocolate Biscuit Bars, Chocolate Confectionery, Chocolate Spread, Confectionery. & Other Exclusions, Cooking Oils, Crisps, Dry Meat Substitutes, Dry Pasta, Dry Pulses and Cereal, Ethnic Ingredients, Everyday Treats, Flour, Frozen Confectionery, Gum Confectionery, Herbal Tea, Herbs and Spices, Home Baking, Honey, Instant Coffee, Lards and Compounds, Liquid and Ground Coffee and Beans, Mincemeat (Sweet), Mustard, Packet Stuffing, Peanut Butter, Pickles Chutneys & Relish, Powder Desserts & Custard, Preserves, RTS. Custard, Ready To Use Icing, RTS Desserts Long Life, Salt, Savoury Snacks, Sour and Speciality Pickles, Special Treats, Suet, Sugar, Sugar Confectionery, Sweet and Savoury Mixes, Syrup & Treacle, Table Sauces, Table and Quick Set Jellies, Tea, Vinegar.
- 5) *Fruit and vegetables*: Ambient Olives, Ambient Rice and Savoury Noodles, Ambient Salad Accompaniment, Baked Bean, Bitter Lemon, Canned Fish, Canned Hot Meats, Canned Salads, Canned Vegetables, Chilled Fruit Juice and Drink, Chilled Olives, Chilled Prepared Fruit and Veg, Chilled Prepared Salad, Chilled Rice, Chilled Salad Accompaniment, Chilled Vegetarian, Cous Cous, Frozen Potato Products, Frozen Vegetables, Frozen Vegetarian Prods, Fruit, Instant Mashed Potato, Nuts, Prepared Peas & Beans, Tinned Fruit, Tomato Products, Total Fruit Squash, Vegetable.
- 6) *Household*: Air Fresheners, Anti-Diarrhoeals, Antiseptics & Liq. Disinfectant, Bath Additives, Batteries, Bin Liners, Bleaches & Lavatory Cleaners, Body Sprays, Carpet Cleaners/Stain Removers, Cat Litter, Cat and Dog Treats, Cleaning Accessories, Cold Sore Treatment, Cold Treatments, Conditioners and Creme Rinses, Contact Lens Cleaners, Cotton Wool, Cough Liquids, Cough Lozenges, Decongestants, Dental Floss or Sticks, Dentifrice, Denture Cleaners/Fixture, Deodorants, Depilatories, Dog Food, Electric Light Bulbs, Eye Care, Fabric Conditioners, Facial Tissues, First Aid Dressings, Foot Preparations, Furniture Polish, Hair Colourants, Hairsprays, Hand Wash Products, Hayfever Remedies, Home Perms, Household Cleaners, Household Food Wraps, Household Insecticides, Incontinence Products, Indigestion Remedies, Kitchen Towels, Laxatives, Liquid Soap, Machine Wash Products, Mens Hairsprays, Mens Mass Fragrances, Mens Skincare, Moist

Wipes, Mouthwashes, Oral Analgesics, Oral Lesion/teething, Pot Pourri and Scented Candles and Oils, Razor Blades, Sanpro, Shampoo, Shaving Soaps, Shoe Care Products, Skincare, Sleeping Aids, Sun Preparations, Talcum Powder, Toilet Soap inc. Mens, Toilet Tissues, Topical Analgesics, Topical Antiseptics, Total Cat Food inc. Bulk, Total Dry Dog Food, Total Male and Female Styling, Total Toothbrushes, Upset Stomach Remedies, Vitamin and Mineral supplements, Wash Additives, Washing Up Products.

- 7) *Meat*: Ambient Cooking Sauces, Ambient Dips, Ambient Pastes and Spreads, Ambient Sandwich Fillers, Ambient Soup, Canned Pasta Products, Chilled Black and White Pudding, Chilled Burgers and Grills, Chilled Cooking Sauces, Chilled Dips, Chilled Gravy and Stock, Chilled Pate and Paste and Spread, Chilled Prepared Fish, Chilled Processed Poultry, Chilled Ready Meals, Chilled Sausage Meat, Chilled Frankfurter/Continental Sausages, Chilled Sandwich Fillers, Cold Canned Meats, Complete Dry/Ambient Meals, Cooked Meats, Cooked Poultry, Fresh Bacon Joint, Fresh Bacon Rashers, Fresh Bacon Steaks, Fresh Beef, Fresh Flavoured Meats, Fresh Lamb, Fresh Other Meat & Offal, Fresh Pasta, Fresh Pork, Fresh Poultry, Fresh Sausages, Fresh Soup, Frozen Bacon, Frozen Beef, Frozen Cooked Poultry, Frozen Fish, Frozen Flavoured Meats, Frozen Lamb, Frozen Meat Products, Frozen Other Meat & Offal, Frozen Pizzas, Frozen Pork, Frozen Poultry, Frozen Processed Poultry, Frozen Ready Meals, Frozen Sausage Meat, Frozen Sausages, Hens Eggs, Instant Hot Snacks, Loose Fresh Meat & Pastry, Meat Extract, Other Chilled Convenience, Other Frozen Foods, P/P Fresh Meat and Veg and Pastry, Packet Soup, Shellfish, Wet or Smoked Fish.
- 8) *Milk*: Total Milk.

APPENDIX B: CONSTRUCTION & REPRESENTATIVENESS OF ESTIMATION SAMPLE

TABLE B1—COMPARISON OF SAMPLE MOMENTS

| Sample:  | Full        |        | Estimation |        | Validation |        | Great Britain |
|--|-------------|--------|------------|--------|------------|--------|---------------|
|  | Mean        | St Dev | Mean       | St Dev | Mean       | St Dev | Mean          |
| Household-level statistics:                    |             |        |            |        |            |        |               |
| Number of Adults                               | 2.0         | 0.8    | 2.0        | 0.8    | 2.0        | 0.8    | NA            |
| Number of Children                             | 0.8         | 1.1    | 0.8        | 1.1    | 0.7        | 1.0    | NA            |
| Household size                                 | 2.8         | 1.4    | 2.8        | 1.3    | 2.7        | 1.3    | 2.4           |
| Characteristics of Household Reference Person: |             |        |            |        |            |        |               |
| Home owner (0/1)                               | 0.74        | 0.41   | 0.78       | 0.41   | 0.76       | 0.43   | 0.69          |
| Age  | 45.5        | 16.0   | 46.9       | 15.8   | 47.0       | 16.0   | 49.2          |
| Retired (0/1)                                  | 0.17        | 0.37   | 0.17       | 0.38   | 0.18       | 0.38   | 0.17          |
| Employed (0/1)                                 | 0.68        | 0.47   | 0.69       | 0.46   | 0.67       | 0.47   | 0.68          |
| Unemployed (0/1)                               | 0.03        | 0.16   | 0.03       | 0.15   | 0.03       | 0.16   | 0.03          |
| Observations:                                  | 26,191×67.6 |        | 2000×3     |        | 2000×3     |        | —             |

*Notes:* The Household Reference Person is a senior member of the household identified using criteria used for the 2001 census in Great Britain. All figures in the column marked Great Britain are for Great Britain from the 2001 Census with the following exceptions: (i) the figure for home ownership is from GB Housing Statistics, rather than the Census, and (ii) the figures for Retired, Employed, and Unemployed status in the last column are for England & Wales only as Scotland does not report this breakdown for the Household Reference Person (when Scotland is eliminated from the Full Sample, the Estimation Sample and the Validation Sample, it does not change the moments reported in the table for these variables).

We implement the sample selection by drawing a week at random for each consumer to represent his third (and final) week in the estimation sample (this must be drawn from outside his first two quarters in the sample). To obtain the second and first weeks in the estimation sample we use the weeks that are one quarter-year and two quarter-years before the third week. When these exact weeks are not available we substitute the most recent available week (that is at least one quarter or two quarters before the final week). We drop consumers for whom three weeks cannot be obtained using this method because they do not participate long enough to be in the data for three successive quarters, which results in a loss of 23 percent of the initial sample of 26,191. We then draw 2000 of the remaining consumers at random to form an estimation sample of 6000 consumer-weeks.

Sample selection problems could arise either because the TNS sample is not representative for the UK population or because the subsample we select is not representative for the full TNS sample. Regarding the latter issue, note that we select consumers almost randomly subject to the constraint that they are in the sample long enough so that we observe each consumer in 3 different quarters. Regarding the former issue, TNS claims to survey a representative sample of consumers and has a commercial interest in making the sample representative.<sup>1</sup> Nevertheless, we analyze explicitly whether the sample is representative by comparing demographics across our sample, the full TNS sample and census data. In Table B1 Full Sample refers to the consumers in the raw sample, Estimation Sample refers to the 2000 consumers selected for estimation, and Validation Sample refers to the 2000 consumers used in the out-of-sample analysis in Section IV. A comparison of sample moments shows that they are similar. The column Great Britain refers to data from 2001 census and allows comparison between the TNS sample means and those of the population.

<sup>1</sup>The commercial value of the data in the form of market analysis for firms requires a representative sample. Kantar describes the panel as a “purchase panel consisting of 30,000 demographically representative households in GB” (<http://www.kantarworldpanel.com/en/Consumer-Panels-/alcohol>”, retrieved 10/7/2016).

C1. Baseline Price Indices

The prices used in the model are computed at category-week-store-demographic group level for categories  $k = 1, \dots, 8$  using the full sample of transactions in the TNS data. (See below for a description of the demographic groups).

In data there are two levels of aggregation below category  $k$ . First, in each category  $k$  (e.g. “Household Goods”), there is a set of narrowly-defined product groups  $g$  (e.g. “Shampoo”) listed in Appendix A. We drop some minor product groups that are not sold by all firms, which leaves 183 (out of 268) product groups that account for 96% of consumer expenditure. We define this set of product groups  $\mathcal{G}_k$  for each  $k$ .

Second, within each product group  $g \in \mathcal{G}_k$  there is a set of products  $h$ , each of which is a unique product and pack size (e.g. “Herbal Essences Fresh Balance Shampoo 200ml” is a product in the “Shampoo” group). Products  $h$  are numerous and there is a tail of products with low volume. For each firm  $f$  we select products  $h$  that appear in the data at least once in each year (2002 - 2005) and in more than six quarterly periods. This yields a set of products,  $\mathcal{H}_{fg}$ , for each firm  $f$  and product group  $g$ . For each store  $j$  product  $h$  and week  $t$  we compute price  $p_{jht}$  as the median price of product  $h$  for week  $t$  for stores operated by store  $j$ ’s firm  $f(j)$ . As noted in Section I the predominant pricing practice is national pricing, in which firms do not vary prices depending on the location of their stores. In cases where there are no observed prices for a particular week we impute the price using the median price for the quarter-year in which week  $t$  falls. We obtain 13 firm-level prices for each  $t$  and  $h$ : one for each of the following: ASDA, Morrison, Sainsbury, Tesco, M&S, Waitrose, Aldi, Lidl, Netto, Iceland, Co-op, and Somerfield, and smaller chains.

The aggregation to category  $k$  level thus proceeds in two stages: (i) from product  $h$  to product group  $g$  and (ii) from product group  $g$  to category  $k$ . In each of these stages we weight the prices to reflect their importance using information from the transactions data.

To allow for taste variation at an intra-category level we compute weights separately for the eight demographic types  $m = 1, \dots, 8$  which are combinations of social class and household size categories. The TNS household characteristics data has six social class levels (1, ..., 6) based on occupational group. These social class indicators are used widely in United Kingdom as a measure of socioeconomic status. A lower number on this scale has a higher average household income. We combine social class level 1 and 2, and likewise 5 and 6, as there are relatively few households in these groups, which yields four social class categories. For each of these we divide households into two size groups—small (one or two people) and large (more than two people)—which yields the eight demographic types.

In the first stage of aggregation the product group  $g$  price in store  $j$  for week  $t$  and demographic group  $m$  is given by  $p_{jgt}^m = \sum_{h \in \mathcal{H}_{gf(j)}} w_{hf(j)}^m p_{jht}$  where  $w_{hf(j)}^m$  are volume weights. We use volume weights at this stage since there is a common volume unit for products within each  $g$  (e.g. volumes in “Shampoo” are in ml). If each product were sold in each firm then we could proceed using volume weights  $w_{hf}^m = Q_h^m / Q_g^m$  where  $Q_h^m$  is the total volume of product  $h$  sold to demographic group  $m$  over the three year period and  $Q_g^m$  is the total volume sold in product group  $g$  to demographic group  $m$  over the three year period. However, each product  $h$  is not sold by all firms so we instead compute  $\tilde{w}_{hf}^m = Q_h^m / Q_{g|h}^m$  where  $Q_{g|h}^m$  is the volume sold in product group  $g$  to demographic group  $m$  by firms selling product  $h$  and let  $w_{hf}^m = \tilde{w}_{hf}^m / \sum_{h \in \mathcal{H}_{gf}} \tilde{w}_{hf}^m$  in order to ensure that the weights add up to one for any firm (i.e.  $\sum_{h \in \mathcal{H}_{gf}} w_{hf}^m = 1$  for any  $f$ ). This weights products using information that is not specific to firm  $f$  for products that are sold by more than one firm and uses firm  $f$  specific information otherwise.

In the second stage of aggregation we obtain the category price  $p_{jkt}^m$  using is a revenue-weighted average of product group price ratios  $p_{jgt}^m / p_{bg}^m$  (where  $p_{bg}^m$  is an arbitrary base price):

$$(C1) \quad p_{jkt}^m = \sum_{g \in \mathcal{G}_k} \omega_g^m \left( \frac{p_{jgt}^m}{p_{bg}^m} \right).$$

The weights  $\omega_g^m$  are the total expenditure share (over the three year period) of each product group  $g$  for demographic type  $m$  (where  $\sum_{g \in \mathcal{G}_k} \omega_g^m = 1$  for each  $m$ ). The weights are constant across stores and over time. Following common practice in price index construction (see for example Chapter 2 in ONS(2014)<sup>2</sup>) we (i) use sales rather than volume weights at this upper level of aggregation because the different product groups are often in different units, and (ii) use price ratios in (C1) to ensure that  $p_{jkt}^m$  is independent of the units chosen within any product group. We set the arbitrary base price  $p_{bg}^m$  in the price ratio to be the price in the first week ( $t = 1$ ) in ASDA stores.

### *C2. Individual Price Indices*

The individual price indices used in subsection V.F differ in the second stage of aggregation by using an individual-specific weighting term—instead of a demographic group weighting term—to aggregate from product group ( $g$ ) to category ( $k$ ) level. (We do the individual weighting at the product group  $g$  level but not the individual product  $h$  level because many individual products such as “private labels” are firm-specific and an individual consumer typically only visits a subset of the firms in the data). The category price  $p_{jkt}^i$  for individual consumer  $i$  is a budget share-weighted average of price ratios  $p_{jgt}^m/p_{bg}^m$  at product group level (where  $p_{jgt}^m$  and  $p_{bg}^m$  are as defined above for the baseline price indices):

$$(C2) \quad p_{jkt}^m = \sum_{g \in \mathcal{G}_k} \omega_g^i \left( \frac{p_{jgt}^m}{p_{bg}^m} \right)$$

where weights  $\omega_g^i$  are now the total expenditure share (over the three year period) of each product group  $g$  by consumer  $i$  and satisfy  $\sum_{g \in \mathcal{G}_k} \omega_g^i = 1$  for each  $i$ . The weights are constant across stores and over time.

<sup>2</sup>Office for National Statistics (2014) “Consumer Price Indices Technical Manual”, available at <http://www.ons.gov.uk>

APPENDIX D: A GENERAL MULTI-STORE-CATEGORY MODEL

In this Appendix we show how the model we estimate can be derived from a more general framework of multi-store and multi-category demand. At the most general level, the consumer chooses for each category  $k \in \{1, \dots, K\}$  in every store  $j \in \{1, \dots, J\}$ , how much quantity  $q_{jk}$  to purchase, subject to his budget constraint:

$$\begin{aligned} \max_{\mathbf{q} \geq 0} \quad & V(\mathbf{q}, \boldsymbol{\theta}, \mathbf{X}) \\ \text{s.t.} \quad & \mathbf{p}'\mathbf{q} \leq y, \end{aligned}$$

where  $\mathbf{q} = (q_{11}, \dots, q_{JK}, q_0)$  denotes the quantity vector for all store/category combinations, and the outside option is  $q_0$ . The price vector  $\mathbf{p} = (p_{11}, \dots, p_{JK}, 1)$  is defined analogously (the price of the outside good is normalized to one).  $\boldsymbol{\theta}$  is a vector of parameters to be estimated and  $\mathbf{X}$  a vector of observable store, category and consumer characteristics.  $y$  denotes the consumer's income.

In this setting corner solutions in quantity are likely to arise and they can originate from two sources. Either the consumer does not visit a particular store and hence cannot purchase any positive quantity there. Or the consumer might visit the store, but decides not purchase any quantity in a specific category. Dealing with the choice over  $J \times K$  quantities with possible corner solutions for many of the quantities makes this a difficult demand system to estimate and we hence impose a set of restrictions based on the data patterns described in Section I of the paper. Specifically, we assume that the cost of visiting more than two stores is prohibitively high, so that no consumer wishes to visit a third store in a given week and that consumers only purchase at one store within a given category.

With these restrictions on the utility function, we can re-write the optimization problem in the following way:

$$\begin{aligned} \max_c \max_{\mathbf{d}} \max_{\mathbf{q} \geq 0} \quad & V(c, \mathbf{d}, \mathbf{q}, \boldsymbol{\theta}, \mathbf{X}) \\ \text{s.t.} \quad & \mathbf{p}_{\mathbf{d}}'\mathbf{q} + q_0 \leq y \end{aligned}$$

This formulation allows us to break up the problem into a discrete choice between (pairs of, or single) stores  $c$ , a discrete choice of store for each category  $\mathbf{d}$  and a continuous quantity choice in each category  $\mathbf{q}$ . To derive equation (1) in Section II of the paper this formulation also assumes utility is additively separable in the variable utility derived from purchasing a specific basket of goods and shopping costs and that variable utility is linear in the outside good. Substituting the budget constraint into the variable utility function for the quantity of the outside option  $q_0$  yields (1).

In this Appendix we discuss the assumptions under which we can aggregate demand from product to category level, and show how this can underpin the utility function we use in the paper. Our derivation follows closely the established literature on aggregation. We adopt the approach of using separability restrictions on preferences (see Gorman (1953) as opposed to that of assuming colinear prices (Hicks (1946)); see Deaton and Muellbauer (1980) for a general treatment.<sup>3</sup>

Let us first define some notation. Let there be  $H$  products and let the quantities bought by a consumer be  $\mathbf{x} = (x_1, x_2, x_3, \dots, x_H)$ . These can be grouped by category using  $\mathbf{x}_k$  so that we may write  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_k, \dots, \mathbf{x}_K)$ . In a similar way let product level prices be  $\mathbf{p} = (p_1, p_2, p_3, \dots, p_H) = (\mathbf{p}_1, \dots, \mathbf{p}_k, \dots, \mathbf{p}_K)$ . To distinguish category level from product level prices we use  $\rho_k$  to denote the category level price index (note this deviates from the notation in the main text). We denote category aggregate quantities  $(q_1, \dots, q_K)$ . Let  $\mathbf{x}_{-k}$  denote a consumption vector for products not in category  $k$ . Weak separability for category  $k$  requires  $(\mathbf{x}_k^1, \mathbf{x}_{-k}) \succeq (\mathbf{x}_k^0, \mathbf{x}_{-k}) \Rightarrow (\mathbf{x}_k^1, \mathbf{x}_{-k}^*) \succeq (\mathbf{x}_k^0, \mathbf{x}_{-k}^*) \forall \mathbf{x}_{-k}^*$  i.e. the quantities an agent consumes of products in other categories ( $\mathbf{x}_{-k}^*$ ) does not change the preferences a consumer has between any two bundles in category  $k$  (here,  $\mathbf{x}_k^1$  and  $\mathbf{x}_k^0$ ). This in turn implies that the consumer's problem may be written:

$$(E1) \quad \max_{\mathbf{x}} u = U[v_k(\mathbf{x}_k), \mathbf{x}_{-k}] \text{ subject to } y = \mathbf{p}\mathbf{x}$$

where  $v_k(\mathbf{x}_k)$  is a category-specific utility function for category  $k$  and  $y$  is the consumer's overall budget.

We can now divide the consumer's problem into two stages: a "first stage" inter-category budget allocation decision in which budget  $y_k$  is allocated to category  $k$  and a number of independent "second stage" problems in which the utility  $u_k = v_k(\mathbf{x}_k)$  from category  $k$  is maximized given the budget  $y_k$ . The indirect utility for the second stage problem is

$$\psi_k(y_k, \mathbf{p}_k) = \max_{\mathbf{x}_k} v_k(\mathbf{x}_k) \text{ subject to } \mathbf{p}_k \mathbf{x}_k = y_k.$$

The first stage decision of how much budget to allocate to category  $k$  can be characterized as a decision of how much category-specific utility  $u_k$  to enjoy, i.e. if category  $k$  is weakly separable (without saying anything about the other categories) we can rewrite (E1) as

$$(E2) \quad \max_{u_k, \mathbf{x}_{-k}} U[u_k, \mathbf{x}_{-k}] \text{ subject to } y = e_k(u_k, \mathbf{p}_k) + \mathbf{p}_{-k} \mathbf{x}_{-k}$$

where the category level expenditure function (dual to the indirect utility function ) is substituted in place of the category  $k$  budget.

If the agent's preferences over products in category  $k$  are homothetic then we can write the category specific indirect utility

$$(E3) \quad \psi_k(y_k, \mathbf{p}_k) = y_k / \rho_k(\mathbf{p}_k)$$

where  $\rho_k(\mathbf{p}_k)$  is the lower-stage price index which must be homogeneous of degree one. From this expression it follows that  $u_k = e_k(u_k, \mathbf{p}_k) / \rho_k(\mathbf{p}_k)$  and hence (by rearranging) the amount budgeted is  $e_k = u_k \rho_k(\mathbf{p}_k)$  which allows us to replace the expenditure function inside the budget constraint expression in (E2) above so that the "first stage" decision can be rewritten

$$\max_{u_k, \mathbf{x}_{-k}} U[u_k, \mathbf{x}_{-k}] \text{ subject to } y = u_k \rho_k(\mathbf{p}_k) + \mathbf{p}_{-k} \mathbf{x}_{-k}.$$

From this it follows that  $u_k$  plays the role of a category-level quantity aggregate and  $\rho_k(\mathbf{p}_k)$  as a category level price index. Using (E3), the category quantity  $q_k$  is obtained by dividing category expenditure  $y_k$

<sup>3</sup>Deaton A. and J. Muellbauer (1980) *Economics and Consumer Behavior*, Cambridge University Press; Gorman, W. (1953) "Community preference fields", *Econometrica*, 21, 63-80; Hicks, J. (1946) *Value and Capital*, 2nd ed., Oxford: Clarendon Press.

by the price index  $\rho_k(\mathbf{p}_k)$ .

We can now derive the utility function we use in the paper. The derivation above extends easily to the case of weak separability between all categories. Under this assumption, we can write overall utility as being composed of category-level sub-utilities.

$$u = U[v_1(\mathbf{x}_1), \dots, v_k(\mathbf{x}_k), \dots]$$

where  $\mathbf{x}_k$  denotes the vector of quantities of goods within category  $k$  and  $v_k()$  denotes the sub-utility function for category  $k$ . Under the additional assumption of homotheticity (within each category), the indirect utility in each category is given by  $y_k/\rho_k(\mathbf{p}_k)$ , where  $y_k$  is category-level expenditure and  $\rho_k(\mathbf{p}_k)$  is a price index.  $y_k/\rho_k(\mathbf{p}_k)$  can be interpreted as a category-level quantity index.<sup>4</sup> We can now think of the utility-function across categories as

$$(E4) \quad u = U[v_1(\mathbf{x}_1), \dots, v_k(\mathbf{x}_k), \dots] = U[q_1, \dots, q_k, \dots] = \boldsymbol{\mu}'_{\mathbf{d}}\mathbf{q} - 0.5\mathbf{q}'\boldsymbol{\Lambda}\mathbf{q} + \alpha q_0$$

This function defines variable utility across all categories (and the outside option). Substituting the budget constraint for  $q_0$  and adding the shopping cost term, we obtain equation (2) in the main text.

<sup>4</sup>Under the assumption of additive (strong) separability between categories, we can replace the assumption that product-level utility is homothetic, with the weaker assumption that product-level utilities have a Generalized Gorman Polar Form.

APPENDIX F: VARIABLES USED IN MOMENTS

- $[\mathbf{Z}_{itcjk}^Q]$ : household  $i$  size  $hz_i$ , time dummies  $\mathbf{T}_t$  (2 years, 3 quarters), price  $p_{tjk}$ , price for categories  $k' \neq k$  (for which we estimate cross effects), log store  $j$  size ( $sz_j$ ), indicator that there are two stores in shopping choice  $c$ ,  $1_{[n(c)=2]}$ , firm dummies (eight of the nine firms in footnote (21)) and a constant.
- $[\mathbf{Z}_{itcjk}^D]$ : as  $\mathbf{Z}_{itcjk}^Q$  but without time dummies  $\mathbf{T}_t$ .
- $[\mathbf{Z}_{itc}^I]$ : distance  $dist_{itc}$ , two-stop shopping indicator  $1_{[n(c)=2]}$ , distance squared  $(dist_{itc})^2$ , interaction of distance and the two-stop shopping indicator  $1_{[n(c)=2]}$ , mean price across categories and stores for shopping choice  $c$  at time  $t$ , and mean price across categories and stores for shopping choice  $c$  at time  $t$  divided by per capita income.

APPENDIX G: LIKELIHOOD FUNCTION

In this Appendix we derive the likelihood function for the model. The observed choice outcome is the triple  $(c, \mathbf{d}, \mathbf{q})$ : a discrete shopping choice  $c$  (up to two stores), a vector  $\mathbf{d} = (d_1, \dots, d_K)$  which indicates the store  $j \in c$  chosen for each category  $k = 1, \dots, K$ , and a  $K$ -vector of continuous choices  $\mathbf{q} = (q_1, \dots, q_K)$  for each category. The likelihood of this choice outcome at parameters  $\boldsymbol{\theta}$  for an individual consumer, written  $L(c, \mathbf{d}, \mathbf{q}|\boldsymbol{\theta})$ , is given by the probability that his unobserved tastes  $(\boldsymbol{\nu}, \varepsilon)$  are in the region that rationalize the choice  $(c, \mathbf{d}, \mathbf{q})$  given the taste density  $f(\boldsymbol{\nu}|\boldsymbol{\theta})$  and the type-1 extreme value distribution for  $\varepsilon$ .

We proceed in four steps. First, in section G.G1 we express the variable utility specification in an alternative but equivalent way that facilitates the derivation of the likelihood. Second, in section G.G2, we derive the set of unobserved tastes  $\boldsymbol{\nu}$  that are consistent with utility maximization given the category choices  $(\mathbf{d}, \mathbf{q})$  at shopping choice  $c$ . Third, in section G.G3, we derive the likelihood of the observed category decisions  $(\mathbf{d}, \mathbf{q})$  treating the shopping choice  $(c)$  as exogenous. Finally, in section G.G4, we derive the joint likelihood of the triple  $(c, \mathbf{d}, \mathbf{q})$  allowing the shopping choice  $(c)$  to be endogenous. Throughout we consider a single consumer-week observation and so to avoid clutter we can suppress  $(i, t)$  subscripts from the notation.

The likelihood is a generalization of likelihoods derived previously for two separate groups of models: (i) those that consider corner solutions for products (or product categories) but do not allow for store choices (Kim et al (2002) and Wales and Woodland (1982)), and (ii) discrete-continuous models that consider only a single continuous choice but do not allow for zero expenditures (Dubin McFadden (1984), Smith (2004)).

*G1. Variable Utility and Category-Store Choices*

In this section we present the variable utility specification in our paper (i.e. the first two terms in (1)) in an alternative but equivalent form that facilitates derivation of the likelihood. Given the specification assumptions of our model, outlined in subsection II.B, we can write the consumer's variable utility from the choice  $(\mathbf{q}, \mathbf{d}, c)$  as

$$\begin{aligned} u(\mathbf{q}, \mathbf{d}, c) &= (\boldsymbol{\mu}_d - \alpha \mathbf{p}_d)' \mathbf{q} - 0.5 \mathbf{q}' \boldsymbol{\Lambda} \mathbf{q} \\ &= \sum_{k=1}^K \sum_{j \in c} [(\mu_{jk} - \alpha p_{jk}) 1(d_k = j)] q_k - 0.5 \sum_{k=1}^K \sum_{k'=1}^K \Lambda_{kk'} q_k q_{k'} \end{aligned}$$

where  $1(d_k = j)$  is an indicator that takes the value 1 if the consumer chooses store  $j$  for category  $k$  and 0 otherwise.

Decomposing the taste coefficient  $\mu_{jk}$  into its deterministic  $(\bar{\mu}_{jk})$  and random  $(\nu_{jk}^\mu)$  components we have

$$(G1) \quad \sum_{k=1}^K \sum_{j \in c} [(\bar{\mu}_{jk} + \nu_{jk}^\mu - \alpha p_{jk}) 1(d_k = j)] q_k - 0.5 \sum_{k=1}^K \sum_{k'=1}^K \Lambda_{kk'} q_k q_{k'}$$

Recall that  $\boldsymbol{\alpha}$  is a random term. It is convenient to write the random effects in (G1) terms of a single unobserved taste term for each  $(j, k)$  so we define  $\nu_{jk}^w = \nu_{jk}^\mu - \alpha p_{jk}$  which implies variable utility is given by

$$(G2) \quad \sum_{k=1}^K \sum_{j \in c} [(\bar{\mu}_{jk} + \nu_{jk}^w) 1(d_k = j)] q_k - 0.5 \sum_{k'=1}^K \sum_{k=1}^K \Lambda_{kk'} q_k q_{k'}$$

or, equivalently, in terms of store-category quantities (written  $q_{jk}$ , where  $\sum_{j \in c} q_{jk} = q_k$ ):

$$(G3) \quad \sum_{k=1}^K \sum_{j \in c} [(\bar{\mu}_{jk} + \nu_{jk}^w) q_{jk}] - 0.5 \sum_{k'=1}^K \sum_{k=1}^K \Lambda_{kk'} (\sum_{j \in c} q_{jk}) (\sum_{j \in c} q_{jk'}).$$

Maximization of (G3) with respect to  $q_{jk}$  for all  $k$  and all  $j \in c$  yields the same outcome as maximization of (G2) with respect to  $\mathbf{q} = (q_1, \dots, q_K)$  and  $\mathbf{d} = (d_1, \dots, d_K)$ . Thus when the shopping choice has two stores ( $n(c) = 2$ ) the utility function implies that consumers use one store per category as an outcome of utility maximization (with respect to quantity  $q_{jk}$  in each store  $j \in c$ ) not an outcome of a constraint on the maximization of utility. This in turn implies that the chosen quantities  $q_{jk}$  satisfy the Kuhn-Tucker conditions for the maximization of (G3) with respect to  $q_{jk}$ , subject to  $q_{jk} \geq 0$ , for all  $k$  and all  $j \in c$ . We use these conditions in the next section.

*G2. Unobserved variable utility tastes implied by choice  $(c, \mathbf{d}, \mathbf{q})$*

We now derive the implications of observed choice  $(c, \mathbf{d}, \mathbf{q})$  and utility maximization for unobserved tastes  $\nu_{jk}^w$  for  $j \in c$ .

When demand at store  $j$  for category  $k$  is positive ( $q_{jk} > 0$ ) the first order condition for maximization of equation (G3) with respect to  $q_{jk}$  is satisfied. This implies that the taste shock  $\nu_{jk}^w$  has a unique value  $\bar{\nu}_{jk}^w$  given by

$$(G4) \quad \bar{\nu}_{jk}^w = -\bar{\mu}_{jk} + \Lambda_{kk}(\sum_{j \in c} q_{jk}) + 0.5 \sum_{k' \neq k} \Lambda_{kk'}(\sum_{j \in c} q_{jk'}).$$

If alternatively the consumer's observed quantity choice for category  $k$  at store  $j \in c$  is zero ( $q_{jk} = 0$ ) the first order condition is not satisfied but the Kuhn-Tucker conditions imply the derivative of utility (G3) with respect to quantity  $q_{jk}$  is not positive which gives the following upper limit for the category-store taste shock

$$(G5) \quad \nu_{jk}^w \leq -\bar{\mu}_{jk} + \Lambda_{kk}(\sum_{j \in c} q_{jk}) + 0.5 \sum_{k' \neq k} \Lambda_{kk'}(\sum_{j \in c} q_{jk'}).$$

The conditions (G4) and (G5) together give the following restrictions on unobserved tastes  $\nu_{jk}^w$  for  $j \in c$  that are consistent with the observed choices  $(c, \mathbf{d}, \mathbf{q})$  and utility maximization:

$$\bar{\nu}_{jk}^w = -\bar{\mu}_{jk} + \Lambda_{kk}(\sum_{j \in c} q_{jk}) + 0.5 \sum_{k' \neq k} \Lambda_{kk'}(\sum_{j \in c} q_{jk'}) \quad \text{if } q_{jk} > 0$$

and

$$\nu_{jk}^w \in \mathcal{A}_{jk} \quad \text{if } q_{jk} = 0$$

where  $\mathcal{A}_{jk}$  is the set of values for  $\nu_{jk}^w$  that are consistent with zero demand for  $(j, k)$ , i.e.

$$(G6) \quad \mathcal{A}_{jk} = \{\nu_{jk}^w \mid \nu_{jk}^w \leq (-\bar{\mu}_{jk} + \Lambda_{kk}(\sum_{j \in c} q_{jk}) + 0.5 \sum_{k' \neq k} \Lambda_{kk'}(\sum_{j \in c} q_{jk'}))\}.$$

Thus we have a unique point value for  $(j, k)$  with positive demand and a set  $\mathcal{A}_{jk}$  of values for  $(j, k)$  with zero demand.

*G3. Likelihood when shopping choice  $c$  is exogenous*

We now use the restrictions on tastes derived in the last section to derive a likelihood  $L(\mathbf{q}, \mathbf{d} \mid \boldsymbol{\theta})$  for the observed choice  $(\mathbf{q}, \mathbf{d})$  assuming that the consumer's shopping choice  $c$  is exogenous. Let

$$(G7) \quad \nu_c^w = (\nu_{j_1}^w, \dots, \nu_{j_K}^w)_{j \in c}.$$

Suppose the consumer is observed to have positive demands for a number  $l$  of store-category  $(j, k)$  combinations and let  $\nu_c^{(1)}$  denote the  $l$ -vector of unobserved tastes  $\bar{\nu}_{jk}^w$  for these. Let  $\nu_c^{(2)}$  be taste shocks for remaining  $(j, k)$  combinations, i.e. those with zero demand. The vector of category-store taste shocks  $\nu_c^w$

in (G7) can be written

$$\nu_c^w = (\nu_c^{(1)}, \nu_c^{(2)}).$$

Let the joint density be  $f(\nu_c^{(1)}, \nu_c^{(2)} | \theta)$ . The likelihood that a consumer selects  $(\mathbf{q}, \mathbf{d})$  combines a probability density component for the unobserved taste elements in  $\nu_c^{(1)}$  (which each have a unique value  $\bar{\nu}_{jk}^w$  given utility maximization at parameters  $\theta$ ) and a probability mass component for the unobserved taste elements in  $\nu_c^{(2)}$  (which each have a range of possible values  $\mathcal{A}_{jk}$  given utility maximization at parameters  $\theta$ ). The likelihood is given by integrating the probability density  $f$  over the range of possible values for  $\nu_c^{(2)}$  at the unique value of  $\nu_c^{(1)}$  i.e.

$$(G8) \quad L(\mathbf{q}, \mathbf{d} | \theta) = \int_{\mathcal{A}(c, \mathbf{q}, \mathbf{d})} f(\bar{\nu}_c^{(1)}, \nu_c^{(2)} | \theta) \text{abs}[\mathbf{J}] d\nu_c^{(2)}$$

where  $\bar{\nu}_c^{(1)}$  is the vector of unobserved store-category taste shocks that satisfy the first order condition in (G4) for all  $(j, k)$  with positive demand. Note that  $\text{abs}[\mathbf{J}]$  is the absolute value of the Jacobian for the transformation from  $\bar{\nu}_{jk}^w$  to  $q_{jk}$  for all the errors in the vector  $\bar{\nu}_c^{(1)}$ . (From (G4) the elements in matrix  $\mathbf{J}$  are given by the second order utility terms in the quadratic utility, i.e.:  $\partial \bar{\nu}_{jk}^w / \partial q_{jk'} = \Lambda_{kk'}$ , etc.). Finally  $\mathcal{A}(c, \mathbf{q}, \mathbf{d})$  is the set of values for unobserved tastes  $\nu_c^{(2)}$  that are consistent with utility maximization given choice  $(c, \mathbf{q}, \mathbf{d})$  and is defined using (G6) as follows

$$\mathcal{A}(c, \mathbf{q}, \mathbf{d}) = \times_{(jk) \in \{(jk) | q_{jk} = 0, j \in c\}} \mathcal{A}_{jk}$$

where  $\times$  denotes the Cartesian product of the sets. The likelihood (G8) is identical in form to equation (7) on page 234 in Kim et al. (2002) and equation (9) on page 266 in Wales and Woodland (1982).

#### G4. Likelihood when shopping choice $c$ is endogenous

We now derive the likelihood  $L(c, \mathbf{q}, \mathbf{d} | \theta)$  for the observed shopping triple  $(c, \mathbf{q}, \mathbf{d})$  allowing shopping choice  $c$  to be endogenous. Unlike the treatment in subsection G.G3 we must now consider shopping costs  $\Gamma_c$  and the variable utility from stores  $j \notin c$ . The observed shopping choice  $c$  depends on the full set of consumer tastes  $\nu$  defined as follows

$$(G9) \quad \nu = \{(\nu_{j1}^w, \dots, \nu_{jK}^w)_{j \in \mathcal{J}}, \nu^\Gamma\}.$$

Let the joint density of these be  $f(\nu)$ . The probability of observing shopping choice  $c$  given unobserved tastes  $\nu$  is given by equation 18. Rewriting this in terms of  $\nu$  we have

$$(G10) \quad P_c(\nu | \theta) = \frac{\exp(w(c, \mathbf{p}_c, \nu_c^w) + \Gamma_c(\nu^\Gamma))}{\sum_{c' \in \mathcal{C}} \exp(w(c', \mathbf{p}_{c'}, \nu_{c'}^w) + \Gamma_{c'}(\nu^\Gamma))}$$

where  $\nu_{c'}^w$  are as defined in (G7) for any  $c' \in \mathcal{C}$ .

We use the restrictions derived in section G.G2 to determine the set of values of  $\nu_c^w$  (for the chosen  $c$ ) that are consistent with the observed choice  $(\mathbf{q}, \mathbf{d})$  and the taste parameters. As in section G.G3 we write  $\nu_c^w = (\nu_c^{(1)}, \nu_c^{(2)})$  where  $\nu_c^{(1)}$  denotes the vector of unobserved tastes  $\nu_{jk}^w$  for  $(j, k)$  combinations (for  $j \in c$ ) with positive demand and let  $\nu_c^{(2)}$  be taste shocks for  $(j, k)$  combinations (for  $j \in c$ ) with zero demand. As well as  $\nu_c^w$  equation (G9) includes  $\nu_{jk}^w$  for  $j \notin c$  and shopping cost shocks  $\nu^\Gamma$ . We group these together as  $\nu^{(3)} = ((\nu_{jk}^w)_{\forall jk | j \notin c}, \nu^\Gamma)$ . The full vector of taste shocks  $\nu$  in (G9) is therefore

$$(G11) \quad \nu = (\nu^{(1)}, \nu^{(2)}, \nu^{(3)}).$$

Let the joint density of these be  $f(\nu^{(1)}, \nu^{(2)}, \nu^{(3)} | \theta)$ .

The joint (discrete-continuous) likelihood that the consumer selects shopping choice  $c$  and category choices  $(\mathbf{q}, \mathbf{d})$  is given by integrating  $P_c(\boldsymbol{\nu}|\boldsymbol{\theta})f(\boldsymbol{\nu})$  over the range of  $\boldsymbol{\nu}$  that are consistent with the consumer making a category choice  $(\mathbf{q}, \mathbf{d})$  at  $c$ . This implies the restrictions for  $\boldsymbol{\nu}^{(1)}$  and  $\boldsymbol{\nu}^{(2)}$  that we derived in subsection G.G3. The likelihood is therefore

$$(G12) \quad L(c, \mathbf{q}, \mathbf{d}|\boldsymbol{\theta}) \\ = \int_{\boldsymbol{\nu}^{(3)}} \int_{\boldsymbol{\nu}^{(2)} \in \mathcal{A}(c, \mathbf{q}, \mathbf{d})} P_c(\bar{\boldsymbol{\nu}}^{(1)}, \boldsymbol{\nu}^{(2)}, \boldsymbol{\nu}^{(3)}|\boldsymbol{\theta})f(\bar{\boldsymbol{\nu}}^{(1)}, \boldsymbol{\nu}^{(2)}, \boldsymbol{\nu}^{(3)}|\boldsymbol{\theta})\text{abs}[\mathbf{J}]d\boldsymbol{\nu}^{(2)}d\boldsymbol{\nu}^{(3)}$$

where  $\bar{\boldsymbol{\nu}}^{(1)}$ ,  $\text{abs}[\mathbf{J}]$ , and  $\mathcal{A}(c, \mathbf{q}, \mathbf{d})$  are defined as in section G.G3. This likelihood resembles the standard likelihood expression  $\int P_c(\boldsymbol{\nu})f(\boldsymbol{\nu})d\boldsymbol{\nu}$  for a mixed logit model for the probability of discrete choice  $c$  in the sense that we integrate the choice probability over the density for  $\boldsymbol{\nu}$ . The difference is that we do not integrate over all possible values of  $\boldsymbol{\nu}$ : we fix some of them (namely  $\boldsymbol{\nu}^{(1)}$ ) to the unique value ( $\bar{\boldsymbol{\nu}}^{(1)}$ ) that is implied by utility maximization given non-zero observed category demands and we restrict others (namely  $\boldsymbol{\nu}^{(2)}$ ) to the set of values  $\mathcal{A}(c, \mathbf{q}, \mathbf{d})$ . Thus instead of obtaining the probability for a discrete choice  $c$  we obtain the probability expression for the discrete-continuous choice of  $(c, \mathbf{q}, \mathbf{d})$ . This likelihood extends those derived in the discrete-continuous literature (Dubin and McFadden (1984), Haneman (1984), Smith (2004)) to the case of multiple-continuous demands with corner solutions.

## H1. Profit maximization in terms of product prices

This subsection demonstrates that the first-order condition (50) can be derived from the assumption of profit maximization at the level of product prices. Let  $p_{fh}$  denote the price of product  $h$  in firm  $f$ , where each product belongs to some category  $k$ .<sup>5</sup> We express this as  $h \in k$ . The profit of firm  $f$  is  $\pi_f(\bar{\mathbf{p}})$ , where  $\bar{\mathbf{p}} = (p_{fh})_{\forall f, \forall h}$ , the vector of all product prices in all firms.

The usual first-order conditions for profit maximization by firm  $f$  are that

$$(H1) \quad \frac{\partial \pi_f(\bar{\mathbf{p}})}{\partial p_{fh}} = 0 \quad \text{for all } h.$$

Suppose that we can aggregate the firm's demand to the category level  $Q_{fk}$  and to write it as a function of category price indices  $p_{fk}$ . Category price indices  $p_{fk}$  are functions of the product prices  $p_{fh}$  so that we can write the function  $p_{fk}(\bar{\mathbf{p}}_{fk})$  where  $\bar{\mathbf{p}}_{fk} = (p_{fh})_{h \in k}$  is the subvector of  $\bar{\mathbf{p}}$  containing only the prices of products  $h \in k$  owned by  $f$ . Then profit can be written

$$(H2) \quad \pi_f(\bar{\mathbf{p}}) = \sum_{k=1}^K Q_{fk}(\mathbf{p})(p_{fk} - mc_{fk}),$$

where  $\mathbf{p} = (p_{fk})_{\forall f, \forall k}$  is the vector of category-specific price indices and  $mc_{fk}$  is the marginal cost. (To simplify the notation, we assume  $\chi_f = 1$  in this discussion. See Section 6.1.) Using (H1) and (H2) we arrive at a first-order condition in terms of the category price index  $p_{fk}$ :

$$(H3) \quad \begin{aligned} 0 &= \sum_{h \in k} \frac{\partial \pi_f(\bar{\mathbf{p}})}{\partial p_{fh}} \\ &= \sum_{h \in k} \frac{\partial}{\partial p_{fk}} \left[ \sum_{k'=1}^K Q_{fk'}(\mathbf{p})(p_{fk'} - mc_{fk'}) \right] \frac{\partial p_{fk}(\bar{\mathbf{p}}_{fk})}{\partial p_{fh}} \\ &= \left[ Q_{fk}(\mathbf{p}) + \sum_{k'=1}^K \frac{\partial Q_{fk'}(\mathbf{p})}{\partial p_{fk}} (p_{fk'} - mc_{fk'}) \right] \sum_{h \in k} \frac{\partial p_{fk}(\bar{\mathbf{p}}_{fk})}{\partial p_{fh}} \\ &= Q_{fk}(\mathbf{p}) + \sum_{k'=1}^K \frac{\partial Q_{fk'}(\mathbf{p})}{\partial p_{fk}} (p_{fk'} - mc_{fk'}) \end{aligned}$$

where the last line follows because  $\sum_{h \in k} \frac{\partial p_{fk}(\bar{\mathbf{p}}_{fk})}{\partial p_{fh}} \neq 0$ . Reintroducing  $\chi_f$ , dividing by  $\partial Q_{fk}/\partial p_{fk}$  and  $p_{fk}$ , and rearranging, we get (50).

## H2. Consumer Group Specific Price Indices

This subsection demonstrates that the first-order condition (50) holds when we allow price indices to vary across consumer groups to reflect different purchasing patterns in households of different size and social class. For simplicity we use the general case of  $i$ -subscripts which allows for price indices for individual consumers. The price index is a weighted average of product prices (see Appendix C)

$$(H4) \quad p_{ifk} = \sum_{h \in k} w_{ih} p_{fh}$$

where  $\sum_{h \in k} w_{ih} = 1$ . To allow for a common shift to all product prices given by the scalar  $\rho_{fk}$  the price index can be written

$$(H5) \quad p_{ifk} = \sum_{h \in k} w_{ih} (p_{fh} + \rho_{fk})$$

<sup>5</sup>Appendix C discusses the construction of price indices based on store-time specific product prices  $p_{jht}$ . Since we look at profit maximization at the weekly level, we suppress the  $t$  subscript in the current discussion.

where  $\rho_{fk} = 0$  at equilibrium prices and

$$(H6) \quad \sum_{h \in k} \frac{\partial p_{ifk}}{\partial p_{fh}} = \frac{\partial p_{ifk}}{\partial \rho_{fk}} = 1 \text{ for all } i.$$

The  $i$ -specific category demands are  $Q_{ifk}(\mathbf{p}_i)$  where  $\mathbf{p}_i = (p_{ifk})_{\forall f, \forall k}$  is the vector of firm-category price indices. Profit is

$$(H7) \quad \pi_f(\bar{\mathbf{p}}) = \sum_i \sum_{k=1}^K Q_{ifk}(p_{ifk})(p_{ifk} - mc_{fk}).$$

As we saw in the previous subsection, profit maximization implies

$$\begin{aligned} 0 &= \sum_{h \in k} \frac{\partial \pi_f(\bar{\mathbf{p}})}{\partial p_{fh}} \\ &= \sum_{h \in k} \sum_i \frac{\partial}{\partial p_{ifk}} \left[ \sum_{k'=1}^K Q_{ifk'}(\mathbf{p}_i)(p_{ifk'} - mc_{fk'}) \right] \frac{\partial p_{ifk}}{\partial p_{fh}} \\ &= \sum_i \frac{\partial}{\partial p_{ifk}} \left[ \sum_{k'=1}^K Q_{ifk'}(\mathbf{p}_i)(p_{ifk'} - mc_{fk'}) \right] \sum_{h \in k} \frac{\partial p_{ifk}}{\partial p_{fh}} \\ &= \sum_i \left[ Q_{ifk} + \sum_{k'=1}^K \frac{\partial Q_{ifk'}}{\partial p_{ifk}} \frac{\partial p_{ifk}}{\partial \rho_{fk}} (p_{ifk'} - mc_{fk'}) \right] \\ &= Q_{fk} + \sum_{k'=1}^K \sum_i \frac{\partial Q_{ifk'}}{\partial \rho_{fk}} (p_{ifk'} - mc_{fk'}) \end{aligned}$$

where the first line is from (H1), the fourth uses (H6), and  $Q_{fk} = \sum_i Q_{ifk}$ . Dividing by  $\frac{\partial Q_{fk}}{\partial \rho_{fk}} = \sum_i \frac{\partial Q_{ifk'}}{\partial \rho_{fk}}$  we have

$$Q_{fk} \left( \frac{\partial Q_{fk}}{\partial \rho_{fk}} \right)^{-1} + (p_{fk} - mc_{fk}) + \frac{\sum_{k'} \frac{\partial \pi_{fk'}}{\partial \rho_{fk}}}{\frac{\partial Q_{fk}}{\partial \rho_{fk}}} = 0$$

where  $p_{fk} = \sum_i \tilde{w}_{ifk} p_{ifk}$  in which  $\tilde{w}_{ifk} = \frac{\partial Q_{ifk}}{\partial \rho_{fk}} / \frac{\partial Q_{fk}}{\partial \rho_{fk}}$ . Note that  $\frac{\partial p_{fk}}{\partial \rho_{fk}} = 1$  which allows us to replace  $\frac{\partial Q_{fk}}{\partial \rho_{fk}}$  with  $\frac{\partial Q_{fk}}{\partial p_{fk}}$ . Rearranging, reintroducing  $\chi_f$ , and dividing by  $p_{fk}$ , we get expression (50).

## APPENDIX I: PROFIT MARGIN CALCULATIONS

In this Appendix we explain how we calculate the profit margin figures which are reported in Table 8. We begin with the calculations using firm-level data covering all grocery categories and then discuss the calculations using data specific to the the milk category which uses the same method.

The Competition Commission (CC) reports two profit margin figures that we use to derive profit margin estimates. The first figure is “gross retail margins”  $m_r$  defined as the difference between the retailer’s annual total revenue and its annual total wholesale cost divided by annual revenue (using the supermarkets’ accounts). The CC reports gross retail margins in the range 0.24 – 0.25 depending on firm (CC(2000) Table 8.19). The second figure is “gross manufacturer margins”  $m_m$  defined as the difference between manufacturer revenues and supplier operating costs (excluding labor costs) as a proportion of manufacturer revenues. The CC reports gross manufacturer margins of 0.25 and 0.36 depending on the sample of firms used (CC(2000) Paragraph 11.108 and CC(2008) Appendix 9.3 Paragraph 11).

Let us begin by deriving a lower bound to the profit margins from these external data. To do this we assume double marginalization, i.e. assume that all payments to manufacturers are of the form of a marginal (or “linear”) wholesale price and the retailer optimizes against this price plus its own marginal costs. Under this assumption the manufacturer’s marginal costs are not relevant to the retailer when setting retail prices so that we can ignore the CC’s information on the manufacturer’s margins. If linear prices are used in relations between supermarkets and manufacturers (as double marginalization implies) then the gross retail margin  $m_r$  is equivalent to the retailers margin over wholesale prices. To obtain the lower bound to the profit margin we combine (i) the assumption of double marginalization, with (ii) the assumption that all of the retailer’s labour costs are marginal, and (iii) the lower end of the range of the figures (noted above) from the CC for  $m_r$  (i.e. 0.24). The CC reports that the ratio of labour costs to wholesale price costs is 9:83 (see CC(2000), Paragraph 10.3) which implies labor costs are  $\frac{9}{83}\% = 10.8\%$  of wholesale costs. This implies we should adjust the retail gross margins reported above using the formula  $m = 1 - 1.108(1 - m_r)$  which gives 0.16. This is the lower bound figure presented in Table 8.

Now we derive an upper bound to profit margins using the external data. To do this we assume that there is efficient retail pricing so that the manufacturer’s marginal cost is relevant to the retailer when setting prices. To obtain an upper bound to margins we combine (i) the assumption of efficient pricing, with (ii) the assumption that *none* of labour costs are marginal, and (iii) the *upper* end of the range of figures (noted above) from the CC for both  $m_r$  and  $m_m$ . With assumptions (i) and (ii) the overall vertical profit margin as a proportion of retail prices is given by the formula  $m = m_r + (1 - m_r)m_m$  where  $m$  is the overall margin,  $m_r$  is retail margin and  $m_m$  is the manufacturer’s margin. Assumption (iii) is that we use the higher of the gross margins figures from the CC for both retailers and manufacturers in this formula, i.s.  $m_r = 0.25$  and  $m_m = 0.36$ . Together this gives the upper bound figure of  $m = 0.52$  that appears in Table 8.

In the case of the milk category the CC reports gross retail margins in the range 0.28-0.30 and gross manufacturer margins in the range 0.04-0.05 (see CC (2008) Appendix 9.3, Paragraphs 12 and 15). Using the same method as in the previous two paragraphs these figures imply margin estimates for the milk category ranging from 0.20 (using  $m = 1 - 1.108(1 - m_r)$  for  $m_r = 0.28$ ) to 0.34 (using  $m = m_r + (1 - m_r)m_m$  for  $m_r = 0.30$  and  $m_m = 0.05$ ).

The lower and upper bounds are conservative because it is likely that some intermediate proportion of labour costs is marginal and because where the CC present a range of figures for gross margins we have (under assumption (iii)) selected them to generate the widest bounds.

APPENDIX J: ALTERNATIVE SPECIFICATIONS

TABLE J1—ESTIMATED PARAMETERS: ALTERNATIVE SPECIFICATIONS

|                | Ind $p_{jk}$ |       | Alt $dist_{ij}$ |       | Alt $\Lambda_{kk'}$ |       | Ind $\Lambda_{kk'}$ |       |
|----------------|--------------|-------|-----------------|-------|---------------------|-------|---------------------|-------|
| $\beta_{01}$   | 2.228        | 0.194 | 2.181           | 0.080 | 2.195               | 0.236 | 2.252               | 0.055 |
| $\beta_{02}$   | 1.475        | 0.049 | 1.456           | 0.031 | 1.497               | 0.076 | 1.477               | 0.052 |
| $\beta_{03}$   | 1.129        | 0.040 | 1.125           | 0.014 | 1.000               | 0.023 | 1.215               | 0.032 |
| $\beta_{04}$   | 2.035        | 0.124 | 1.972           | 0.047 | 1.891               | 0.042 | 2.068               | 0.096 |
| $\beta_{05}$   | 2.783        | 0.208 | 2.731           | 0.152 | 2.823               | 0.131 | 2.863               | 0.226 |
| $\beta_{06}$   | 1.113        | 0.033 | 1.088           | 0.016 | 1.181               | 0.038 | 1.129               | 0.026 |
| $\beta_{07}$   | 2.359        | 0.133 | 2.617           | 0.113 | 2.419               | 0.127 | 2.566               | 0.099 |
| $\beta_1$      | 0.469        | 0.003 | 0.437           | 0.004 | 0.449               | 0.004 | 0.415               | 0.004 |
| $\beta_2$      | 0.450        | 0.011 | 0.482           | 0.011 | 0.472               | 0.018 | 0.469               | 0.013 |
| $\sigma_1$     | 0.188        | 0.032 | 0.320           | 0.018 | 0.213               | 0.030 | 0.197               | 0.039 |
| $\sigma_2$     | 0.977        | 0.022 | 1.046           | 0.027 | 0.894               | 0.033 | 0.973               | 0.026 |
| $\sigma_3$     | 0.580        | 0.016 | 0.731           | 0.026 | 0.579               | 0.037 | 0.751               | 0.039 |
| $\sigma_4$     | 1.186        | 0.023 | 1.091           | 0.024 | 1.096               | 0.029 | 1.077               | 0.020 |
| $\sigma_5$     | —            | —     | —               | —     | —                   | —     | 0.225               | 0.045 |
| $\Lambda_{11}$ | 22.254       | 6.828 | 21.116          | 3.305 | 20.153              | 3.019 | 21.542              | 2.248 |
| $\Lambda_{22}$ | 11.851       | 0.655 | 12.093          | 0.375 | 11.840              | 3.499 | 11.996              | 0.746 |
| $\Lambda_{33}$ | 3.919        | 0.286 | 4.035           | 0.182 | 3.601               | 0.198 | 4.330               | 0.218 |
| $\Lambda_{44}$ | 11.485       | 1.087 | 11.276          | 0.412 | 10.214              | 0.420 | 11.651              | 0.866 |
| $\Lambda_{55}$ | 16.851       | 2.558 | 16.808          | 1.164 | 16.554              | 1.060 | 17.219              | 2.049 |
| $\Lambda_{66}$ | 4.315        | 0.204 | 4.348           | 0.144 | 4.489               | 0.241 | 4.369               | 0.204 |
| $\Lambda_{77}$ | 9.105        | 0.902 | 10.677          | 0.674 | 9.127               | 0.792 | 10.094              | 0.653 |
| $\Lambda_{88}$ | 13.633       | 0.392 | 14.427          | 0.568 | 15.572              | 0.775 | 14.395              | 1.407 |
| $\Lambda_{57}$ | 1.879        | 0.121 | 2.032           | 0.081 | —                   | —     | 2.070               | 0.115 |
| $\Lambda_{18}$ | 1.558        | 0.083 | 1.444           | 0.182 | —                   | —     | 1.107               | 0.605 |
| $\Lambda_{23}$ | 0.286        | 2.473 | 0.201           | 0.681 | —                   | —     | 0.382               | 1.269 |
| $\Lambda_{23}$ | 0.394        | 1.709 | 0.870           | 0.998 | —                   | —     | 0.554               | 0.346 |
| $\Lambda_{23}$ | 0.061        | 1.030 | 0.113           | 0.152 | —                   | —     | 0.160               | 0.656 |
| $\Lambda_{14}$ | —            | —     | —               | —     | 0.717               | 0.426 | —                   | —     |
| $\Lambda_{27}$ | —            | —     | —               | —     | -0.033              | 1.213 | —                   | —     |
| $\Lambda_{38}$ | —            | —     | —               | —     | -0.321              | 0.172 | —                   | —     |
| $\lambda_1$    | —            | —     | —               | —     | —                   | —     | 0.241               | 0.052 |
| $\lambda_2$    | —            | —     | —               | —     | —                   | —     | -0.033              | 0.062 |
| $\alpha_1$     | 1.759        | 0.056 | 1.908           | 0.021 | 1.938               | 0.033 | 1.967               | 0.043 |
| $\alpha_2$     | 34.374       | 2.351 | 32.315          | 2.544 | 37.584              | 3.503 | 20.961              | 5.116 |
| $\gamma_{11}$  | 8.477        | 1.135 | 8.762           | 0.849 | 8.660               | 1.250 | 9.206               | 1.237 |
| $\gamma_{12}$  | 0.447        | 0.028 | 0.436           | 0.026 | 0.457               | 0.029 | 0.402               | 0.024 |
| $\gamma_{21}$  | 11.636       | 2.045 | 11.188          | 1.613 | 12.773              | 2.340 | 13.142              | 2.273 |
| $\gamma_{22}$  | -0.386       | 0.031 | 0.385           | 0.029 | 0.422               | 0.032 | 0.366               | 0.028 |

Notes: See notes for Table 3. Specifications described in subsection V.F.