Supermarket choice with multi-store shopping: measuring the effect of format regulation*

Pasquale Schiraldi[†], Stephan Seiler[‡], Howard Smith[§]

September 19, 2011

Abstract

In recent years the growth of the big box supermarket format, with its large stores and product range, has been regulated by UK public policy, with the goal of protecting smaller town-centre stores. Whether this restriction reduces competition—as some have argued—depends on how much consumers substitute between large and small store formats. This paper analyzes these questions using a model in which consumers can combine stores of different formats in a given shopping period. The presence of multi-store shoppers allows an extra source of substitution through the choice of how much to spend in each store, which brings stores that are combined into greater competition with each other. We estimate the model on a dataset of consumer choices and evaluate the effect—on market power and on town centres—of entry regulation by adding a list of stores that firms wanted to open but that the planner rejected.

PRELIMINARY AND INCOMPLETE—DO NOT CIRCULATE, QUOTE OR CITE

*We are grateful to the Webb Medley Fund for support.

[†]London School of Economics

[‡]London School of Economics

[§]Oxford University Economics Department

1 Introduction

Supermarket stores vary widely in the range and quality of products they offer. The "big-box" format—located in large stores outside of town centres—is attractive for its wide product range. Other formats offer narrower product ranges, focusing on a specific quality, price, or geographic location. The emergence of these distinct formats has led to public policy questions. The first is whether big-box retailing has an adverse impact on small stores and consequently the environmental quality of town centres.[1] The second is whether restrictions on development of big box retailers—resulting in a high level of concentration within the format—is enough to confer market power regardless of the presence of outlets from other formats.

In many countries public policy has recognized the importance of store format. In the UK environmental and competition policy authorities take distinct positions. Environmental planning authorities have developed a policy of protecting town centres. To promote this policy restrictions are placed on the development of big box stores (which are generally not in town centre locations). Competition authorities, on the other hand, have claimed that these restrictions reduce competition within the big-box format: they claim that these are the only format suitable for one-stop shopping trips, and, for consumers on such trips, small-store formats are treated as poor substitutes. [2] Until recently these two areas of policy have developed independently but currently there is a debate over how much planning policy should respond to competition as well as environmental policy issues.

The appropriate policy response depends on how readily consumers substitute between different supermarket formats. If consumers substitute mainly between stores within a format then the opening of a big-box store will have a beneficial competi-

[1] We do not comment on whether protection of town centre stores, and town centre vitality, are worthwhile public policy objectives.

[2] As a consequence, mergers between firms in different formats are treated more leniently than mergers in the largest formats.The most dominant big box retailers Tesco, ASDA and Sainsbury were prohibited from merging with Safeway, which also operated big box stores. Inter-format policy in the UK: the merger of ASDA and Netto was permitted, while Tesco and Sainsbury have been allowed to expand market share by acquisition (and new openings) of small-format stores.

tive effect within its format and a limited adverse effect on the smaller formats that operate in town centres. But if consumers substitute readily between formats—if competition is diffuse—the store might have more limited competitive benefits and more adverse effects on town centres. In the latter case, planning restrictions are more easily justified.

A traditional framework for analyzing retail competition is to assume consumers are characterized by one-stop shopping. One-stop shoppers select a single store and stores that have the most similar products and prices are likely to compete the most vigorously. If, on the other hand, consumers engage in multi-stop shopping—combining distinct stores within a shopping period—they are able to substitute demand between these stores. This reduces the market power of stores (especially those for which a high proportion of shoppers are multi-stop shoppers). Furthermore, substitution depends on which stores are *combined*: if consumers like to combine stores of different formats then, a "big box" store might compete strongly with stores of other formats. In practice there is a mix of shopper types and consumers choose between one-stop and multi-stop shopping, depending on the consumer's shopping costs and the opportunities for shopping benefits in the consumer's choice sets. As multi-stop shopping is endogenous in this way, there is a further consequence of the "big-box" format: it may induce a change from multi-stop to one-stop behavior, and if this effect is strong enough a big box retailer—converting multi-stop to one-stop shoppers—may result in adverse effects both for both town-centres and—by removing a source of substitution—for pricing competition between retailers.

This paper evaluates the effects of big box retailing using an empirical model of consumer choice that is specified at the level of the individual consumer and store. Consumers choose one or multiple stores from their local choice set. The attributes of the stores and the product range available determine these choices. We allow some stores to be better in combination than others depending on the stores product range and characteristics. The consumers also make a continuous choice of how much to buy in each store. This is a "discrete-continuous" consumer model, with multiple discrete choices of stores and multiple continuous budget allocations between the chosen stores, for each of a number of grocery categories. The model allows a price increase at a store to have the following two effects on grocery demand at the store: (i) an effect from consumers who continue to visit the store but switch some of their continuous expenditures from the store (to other stores if the consumer is a multi-stop shopper) and (ii) another effect as the price increase reduces the number of consumers visiting the store. The model predicts the shopping behavior of consumers in a unified utility framework.

We estimate the model for the supermarket industry in Great Britain using a dataset of consumer choices and store characteristics. The consumer survey tracks the complete grocery shopping behavior of respondents over time, including the choice of store, the expenditure at the store, the number of product lines available in each type of store, and the locations and demographics of consumers. A second dataset of store characteristics contains store size and location characteristics. These datasets are matched so that the consumer's choice set is known as well as the details of the chosen stores. We allow the benefits of two stop shopping to vary across hypothetical store pairs, so that some store characteristics (such as big box format) may be associated with greater one stop shopping than others. We also allow some store characteristics to interact positively in utility—and others to interact negatively—so that different types of stores are combined to different extents.

We scale the estimated model to the level of Great Britain. We the model to analyze the effect of policy towards store formats. We evaluate counterfactuals in which large format stores are added and removed from the market. We evaluate the effect on town centres of such stores. In particular, we evaluate the effect of adding stores that were rejected by the planning authority, and assess their impact on town centre stores.

The theory of supermarket competition has been studied in a one-stop shopping context by Bliss (1988) and in a two-stop context by Lal and Matutues (1994) and

Klemperer (1992). The latter, particularly related to our paper, studies the incentives for firms to manipulate their store format to minimize the extent of two-stop shopping.

The empirical model of consumer choice that is estimated here is related to Dubin and McFadden (1984) and Smith (2004) but extends these models to allow for multiple discrete choices and several categories of spending. In allowing for pairs of stores to be combined, our paper is related to Genzkow (2007). Smith (2004) uses a single price index for all groceries, this paper allows two-store shoppers to allocate spending between the two stores differently for different categories, and the price index is endogenised to weigh the category prices endogenously.[3]

The paper is also related to the literature on Wal-Mart, which has mainly approached the topic using entry models (Holmes (2011) and Jia (2008)). Unlike these models we use demand side data to evaluate the effect of big-box retailing, and exploit the variation across markets in the presence of these stores.

The rest of the paper is as follows. In section 2 we present some background information and introduces the data. In section 3 we set out the theory model. In section 4 we discuss the estimation of the model. Section 5 presents the estimates and results and section 6 concludes.

2 Policy Background and Data Summary

2.1 Policy Background

Planning Policy In the UK there was a major planning policy change dating from the mid 1990s. From 1996 planning policy was changed resulting in a much more restrictive attitude to the entry of new big box retailing outlets, with the aim of protecting town centres. To open a store requires planning consent and the criteria in the new planning guidelines are: (a) a *sequential test* that asks (for big-box out-of-centre applications) if development at a town centre location was possible instead; (b) a *need* test that asks

[3] A further literature, Wales and Woodland (1983), and other Kuhn-Tucker based utility specifications for dealing with corner problems, is related to this paper

if there is a "need" for the extra floorspace conditional on existing floorspace and town population; and (c) a test for whether the development will adversely affect business in the town centre.[4] The planning policy is much less restrictive for alternative formats. Since the introduction of tighter planning criteria new store formats have changed greatly with a much higher proportion of new stores being of small-store format in town centre locations, and a much smaller proportion being big-box format.

Competition Policy. The supermarket industry has been investigated in a series of reports from the Competition Commission (CC) in 2000, 2006, and 2010. The CC argued that smaller format stores were poor substitutes for big box stores for consumers undertaking a large one-stop shopping trip, who do not consider small stores attractive. However, the CC argued, for consumers doing smaller multiple-stop trips, small and large stores were able to compete. A consequence of the change to planning policy from 1996 was to place a barrier to the entry of new big-box stores, locking in the existing market structure for big-box stores, which is often quite concentrated. For this reason, the CC opposed further mergers of two firms within the big box format, though they allowed other mergers (e.g. the recent Netto/ASDA merger where only one firm was in the big-box format).

2.2 Data

The data comprises a dataset of store characteristics and a survey of consumer choices covering the period Oct 2002-Sept 2005. The store characteristics data (source: IGD) is a complete dataset of all stores open in the period in Great Britain, and gives store size, firm, location (by postcode, which yields an exact grid reference), and the type of location (town centre, out of town, etc). For any consumer location we can use this data to compute choice sets of nearby stores. The consumer data comes from a survey run by TNS which records the daily purchases, and stores, of several thousand consumers over the three years in various locations across Great Britain. For the most important

[4]The criteria for (b) are relatively observable, and (c) can be obtained using our demand model. The criteria for (a) may depend on unobservable factors to do with available sites in the centre.

six firms we know the exact store that the consumer visited, and for the other firms, we know the firm (but not exact store). The consumer data is at individual product level, and we are able to record the number of distinct products sold in firms of alternative size classes. Demographic information on the consumers is also recorded, including location, social class, and household size. We aggregate spending into 6 broad product categories. For each of these categories we compute a price index for each firm and week using the product-level prices observed in the consumer data, and these price indices are computed separately for each of 8 demographic groups to reflect their different tastes. Further details of the data are given in the appendix.

We also use data that indicates whether a store is located in the town centre: using data on retail floorspace density and economic activity indicators, town centres have been defined by the Department of Communities and Local Government for England and Wales.

2.3 Store formats and consumer behavior

The Firms and their Store Formats Table 1 shows the store characteristics of the main firms, including product range and pricing. The firms typically have national pricing policies, that mean that the price of a product sold by a firm is the same regardless of which store or region of Great Britain it is sold in.

The first four firms—ASDA, Morrisons, Sainsbury, and Tesco—operate the "big-box" stores. (ASDA is WalMart's UK subsidiary, though its stores are somewhat smaller than WalMart stores in the US). These stores have floorspace of around 40,000 square feet and stock around 40,000 distinct product lines. Some of these firms also operate smaller store formats, with fewer product lines.

Under the heading "Discounter" we group three very similar firms that specialize in small stores of about 8000 square feet and about 6000 distinct product lines, sold at low prices (Aldi, Lidl and Netto). Iceland specialize in frozen food and Somerfield and Co-op (as well as firms grouped under the heading "Other") are attractive mainly because of a convenient location. Marks & Spencer and Waitrose specialize in high quality food, at higher prices, and of these two firms Waitose has a much more comprehensive set of product lines.

Table 2 indicates the extent of product overlap for stores of the main firms. Firms vary greatly in the extent of product overlap. Only a small proportion of Marks & Spencer products are available elsewhere. The figure for two large stores is about 50%. The presence of private label products means that this figure rarely rises above 50%.

Consumer Shopping Patterns We treat a weekly period as the most natural shopping period to use. This is by far the most commonly observed shopping frequency. Table 3 shows the extent of one-stop and multi-stop shopping in a weekly shopping period. Multi-stop shopping is quite common: 40% of shoppers visit two or more stores in a week. This is concentrated in the first store: 86% of shopping is in the largest weekly store. However a significant amount is in the second store—some 11%. Overall 93% of spending is concentrated in the top two trips. Note that multi-trip shopping is more common than multi-store shopping: 50% of two trip shoppers are visiting the same store twice. Nonetheless, multi-store shopping is some 39% of trips, so it represents a significant portion of shoppers.

Table 4 indicates that some pairings of firms are more popular than others. In particular we see that the big box stores are least likely to be combined with any other store, while the small stores are more likely to be combined with others. Table 5 shows that a pair involving a large and a small store is more likely to be chosen than two large (or two small) stores. [Other tables to show the relationship between of two-stop shopping and expenditure].

Table 6 shows that different firms (regardless of store size) specialize in different product categories. Breaking the grocery demand into six self-explanatory categories—alcohol, chilled, dry, frozen, fresh (but not chilled), and household products—we see for example that M&S is very strong for chilled foods but not household products, Iceland is strong for frozen food, and the Discounters are strong for storable products such as dry groceries and alcohol. [Another table here will show that store size is important in the size of trip].

3 The Consumer Model

Consumer i at time t chooses either one or two stores from the nearest 30 stores. More formally i makes a choice $c = (j, j') \in C_i$ where j and j' are an pair drawn without replacement from a choice set J_i comprising the 30 nearest stores and the addtional option of going to no store. We write $j = 0$ to denote the option of no store. The choice set C_i is all (unordered) pairs drawn without replacement from J_i, which implies $\binom{31}{2} = 465$ choices for each consumer. Consumers vary according to their location and belong to one of the eight demographic groups d defined in the previous section. The consumer selects a quantity q_{jkt} for each chosen store and each of the six broad expenditure categories $k = 1, .., 6$ defined in the previous section, and $q_{jkt} = 0$ for all k if $j = 0$. To reflect the different weights that demographic groups attach to different products within each category k we construct a price index separately for each demographic group $p^{d(i)}_{jkt}$.

Consumer i of demographic type $d(i)$ purchasing $(q_{jkt}, q_{j'kt})$ units of category k at choice $c = (j, j')$ at time t obtains gross utility

$$u_{ickt}(q_{jkt}, q_{j'kt}) = \sum_{j \in (j,j')} \left(\gamma_x x_{jk} + \xi^{d(i)}_{f(j)k} + \nu_{ijkt} \right) q_{jkt} - \sum_{j \in (j,j')} \frac{\gamma_k}{2} q^2_{jkt} - \sigma_{ck} q_{jkt} q_{j'kt} \quad (1)$$

and marginal utility of category k at store j is

$$\frac{\partial u_{ickt}}{\partial q_{jkt}} = \left(\gamma_x x_{jk} + \xi^{d(i)}_{f(j)k} + \nu_{ijkt} \right) - \gamma_k q_{jkt} - \sigma_{ck} q_{j'kt}. \quad (2)$$

$f(j)$ denotes the firm operating store j and x_{jk} is other observable store characteristics (which may vary by k such as number of product lines). $\xi_{f(j)k}$ is a firm-category dummy that allows firms to differ in their mean attractiveness for a given category. Individual i's deviation from this mean is given by ν_{ijkt} which combines a time-constant firm-specific component $\nu^1_{if(j)}$, a time-constant category-specific component ν^2_{ik}, and a time-varying component ν^3_{ikt}:

$$\nu_{ijkt} = \mu_{1k}\nu^1_{if(j)} + \mu_{2k}\upsilon^2_{ik} + \mu_{3k}\upsilon^3_{ikt}$$

so that we have variation across consumers in taste for each category, in taste for each firm, and in their taste for a category-firm combination. The terms $\upsilon^1_{if(j)}$, υ^2_{ik} and υ^3_{ikt} are each drawn from normal distributions $N(0,1)$, and $(\mu_{1k}, \mu_{2k}, \mu_{3k})$ are k-specific scaling terms, allowing taste heterogeneity to differ by category.

The expression (2) for marginal utility of category k at store j declines with the quantities bought at each store, q_j and $q_{j'}$. The parameter γ_k governs the rate at which marginal utility declines in q_{jkt}. $\sigma_{ck} \in [0, \gamma_k]$ is a substitution parameter that governs the rate at which marginal utility declines in the other-store quantity $q_{j'kt}$: if $\sigma_{ck} = 0$ the two stores' products are independent in utility and consumers can gain utility by buying from two stores, whereas if $\sigma_{kc} = \gamma_1$ the two store's products are close substitutes in this category and the consumer will not buy the category at both stores. We specify the substitution term as follows

$$\sigma_{ck} = \gamma_k \exp(\sigma_{1k} + \sigma_2 x_{ck})/(1 + \exp(\sigma_{1k} + \sigma_2 x_{ck}))$$

where x_{ck} is observable variables that measure the closeness of the two store's product lines in category k (e.g. degree to which the two stores' products overlap in category k).

We assume a quasi-linear structure in which there is a further outside good q_0 that enters linearly in the consumer's utility. We normalise the price of the outside good to unity and assume the marginal utility of money α_i is constant for each consumer over time. The total net utility consumer i at time t derives from choice c and quantities

$q_{ct} = \{q_{jkt}, q_{j'kt}\}_{k=1}^{K}$ is as follows

$$\sum_{k=1}^{K} \left[u_{ickt} \left(q_{jkt}, q_{j'kt} \right) - \alpha_i \sum_{j \in (j,j')} p_{jkt}^{d(i)} q_{jk} \right] + \alpha_0 + \beta_i X_{ict} + \mu_4 \varepsilon_{ict}$$

where $\beta_i X_{ict} + \mu_4 \varepsilon_{ict}$ represents the costs of choice c that do not influence the quantities purchased. X_{ict} is a vector of observable characteristics of choice c—including distance from consumer i and interactions of the two store's characteristics—and ε_{ict} is an iid stochastic error with a standard Type 1 Extreme Value distribution; μ_4 is a scaling term. Neither $\beta_i X_{ict}$ or ε_{ict} influence the marginal utility from each category conditional on choice c, through they do influence the choice c. The taste β_{il} on the l'th element of X_{ict} varies by consumer such that

$$\beta_{il} = \beta x_i + \mu_l \nu_{il}$$

where x_i is an observable demographic attribute (such as car ownership where X is distance) and ν_{il} is a random draw. We include in X_{ict} a dummy for whether one or two stores is chosen, and the random coefficient on this allows the cost of shopping in more than one store per week to vary across consumers.

We specify α_i to be a function of consumer income so that high-income households may be less price-sensitive than low income households..

Utility is additively separable in the categories k. The assumption is strong but we consider it reasonable because the categories are broad. It can be relaxed by adding interaction terms between (possibly nested) categories, which brings the cost of extra parameters to estimate.

Utility theory permits two normalizations. First, as only utility differences matter, we can set $\alpha_0 = 0$ as it is common to all choices. Second, as we cannot identify the level of utility, we can set $\mu_4 = 1$, which eliminates this parameter, and rescale all the remaining parameters accordingly (including the γ, ξ in parameters, which enter $u()$ linearly).

Consumer i's indirect utility $w_{ickt}(p_{ct}^{d(i)})$ from category k conditional on choice c given prices $p_{ckt}^{d(i)} = \left\{p_{jkt}^{d(i)}\right\}_{j\in c}$ is given by maximization of utility subject to non-negativity constraints for the quantities of each category at each store:

$$w_{ickt}(p_{ckt}^{d(i)}) = \max_{q_{jkt}, q_{j'kt}} u_{ickt}\left(q_{jkt}, q_{j'kt}\right) - \alpha_i \sum_{j\in(j,j')} p_{jkt}^{d(i)} q_{jkt}$$

$$\text{subject to } q_{jkt} \geq 0 \qquad j = (j, j').$$

Optimal quantities $\tilde{q}_{ijckt} = \tilde{q}_{ijckt}(p_{ckt})$ for $j = (j, j')$ are implied by the Kuhn-Tucker complementary slackness conditions:

$$\left\{\left(\frac{\partial u_{ikt}}{\partial q_{jkt}} - \alpha_i p_{jkt}^{d(i)}\right) \leq 0; \qquad q_{jkt} \geq 0; \qquad q_{jk}\left[\frac{\partial u_{ikt}}{\partial q_{jkt}} - \alpha_i p_{jkt}^{d(i)}\right] = 0\right\} \qquad \text{for } j = (j, j').$$

The model generates both interior solutions where positive quantities are bought and boundary solutions where the quantity is zero (at one or both stores). When a positivel quantity is predicted at store j the condition for optimal q_{jkt} is:

$$\frac{\partial u_{ikt}}{\partial q_{jkt}} - \alpha_i p_{jkt}^{d(i)} = \gamma_x x_{jk} + \xi_{f(j)k}^{d(i)} + \nu_{ijkt} - \gamma_k q_{jkt} - \sigma_{ck} q_{j'kt} - \alpha_i p_{jkt}^{d(i)} = 0 \tag{3}$$

which implies the conditional category demand

$$q_{ijckt} = \frac{1}{\gamma_k}\left(\gamma_x x_{jk} + \xi_{f(j)k}^{d(i)} - \sigma_{ck} q_{j'kt} - \alpha_i p_{jkt}^{d(i)} + \nu_{ijkt}\right). \tag{4}$$

The own-price effect on demand for category k (conditional on store choice) is given by α_i/γ_k, so that variation in γ_k allows own-price effects to vary across categories.

Consumer i visiting store combination c obtains total indirect utility

$$w_{ict}(p_{ct}^{d(i)}) = \sum_k \left[w_{ickt}(p_{ckt}^{d(i)})\right] - \beta^i X_{ict} + \varepsilon_c$$

where $p_{ct}^{d(i)} = \left\{ p_{ckt}^{d(i)} \right\}_{k=1}^{6}$ and choice of c satisfies the condition:

$$i \text{ chooses } c \text{ at time } t \Longleftrightarrow w_{ict}(p_{ct}^{d(i)}) > w_{ic't}(p_{c't}^{d(i)}) \text{ for all } c'.$$

Consumer i's demand in store j for category k at time t is given by:

$$q_{ijkt} = \sum_{c \in c(j)} \tilde{q}_{ijktc}(p_{ct}^{d(i)}).1 \left[w_{ict}(p_{ct}) > w_{ic't}(p_{ct}) \right]$$

where 1[] is an indicator function and $c(j)$ is the set of all choices c in C_i that contain store j.

We rewrite the category-level demand and indirect utility functions conditional on choice c to make explicit their dependence on the consumer's private shocks $\nu_{ickt} = \{\nu_{ijkt}\}_{j \in c}$:

$$\begin{array}{ll} \tilde{q}_{ijckt}(\nu_{ickt}) & \text{store-category demand conditional on } c \\ w_{ickt}(\nu_{ickt}) & \text{indirect utility for category } k \text{ conditional on } c \end{array}$$

Similarly we write the overall indirect utility for choice c conditional on private shocks $\nu_{ict} = \left[\{\nu_{ickt}\}_{k=1}^{6}, \nu_{il}, \right]$ and ε_{ict}:

$$w_{ict}(\nu_{ict}, \varepsilon_{ict}) \quad \text{indirect utility for choice } c.$$

The parameters are summarised as follows:

$$\theta \equiv (\lambda, \xi) \text{ where } \lambda = (\alpha, \beta, \gamma, \mu, \sigma).$$

We assume that ε_{ict} is distributed Type-1 Extreme Value, so that conditional on a particular draw $\nu_{it} = (\nu_{ict})_{c \in C_i}$ for the consumer, the parameters θ imply a probability $P_{ict}(\theta | v_{it})$ of consumer i making the choice c at time t given by the multinomial logit probability:

$$
\begin{aligned}
P_{ict}(\theta|v_{it}) &= \Pr(w_{ict}(v_{ict},\varepsilon_{ict}) > w_{ic't}(v_{ict},\varepsilon_{ict})\ \forall c \in C_i \mid \nu_{it}) \\
&= \frac{\exp\left(\sum_k \left(w_{ickt}(\nu_{ickt}) - \alpha_i \sum_{j\in(j,j')} p_{jkt}^{d(i)}\tilde{q}_{ijckt}(\nu_{ickt})\right) - \beta_i X_{ict}\right)}{\sum_c \exp\left(\sum_k \left(w_{ickt}(\nu_{ickt}) - \alpha_i \sum_{j\in(j,j')} p_{jkt}^{d(i)}\tilde{q}_{ijckt}(\nu_{ickt})\right) - \beta_i X_{ict}\right)}
\end{aligned}
$$

The unconditional probability $P_{ict}(\theta)$ of choosing choice c is given by integrating over the distribution of unobserved consumer tastes $F(v_{it})$:

$$
P_{ict}(\theta) = \Pr\left(w_{ict}(v_{it},\varepsilon_{ict}) > w_{ic't}(v_{it},\varepsilon_{ic't})\ \forall c'\right) \tag{5}
$$

$$
= \int \frac{\exp\left(\sum_k \left(w_{ickt}(\nu_{ickt}) - \alpha_i \sum_{j\in(j,j')} p_{jkt}^{d(i)}\tilde{q}_{ijckt}(\nu_{ickt})\right) - \beta_i X_{ict}\right)}{\sum_{c'} \exp\left(\sum_k \left(w_{ickt}(\nu_{ic'kt}) - \alpha_i \sum_{j\in(j,j')} p_{jkt}^{d(i)}\tilde{q}_{ijc'kt}(\nu_{ickt})\right) - \beta_i X_{ic't}\right)} dF(\nu_{it})
$$

Consumer i's expected demand in store j for category k is given by summing over the predicted demands for the subset $c_i(j)$ of choices in C_i that contain the store j

$$
E(q_{ijkt}) = \sum_{c\in c_i(j)} \int \tilde{q}_{ijckt}(\nu_{ickt}) \Pr(w_{ict}(v_{it}) > w_{ic't}(v_{it})\ \forall c|\ \nu_{it}) dF(\nu_{it}).
$$

4 Estimation

We estimate the model using the method of simulated moments (McFadden (1989)), using two sets of moments respectively based on the model's predictions for: (i) the discrete choice $c = (j, j')$ and (ii) the continuous quantities q_{ijkt} for each j, k and t.

For the first set of moments, based on the discrete choice c, we use the residuals:

$$
d_{ict} - P_{ict}(\theta) \text{ for all } i \text{ and } t, \text{ and all } c \in C_{it}
$$

where $P_{ict}(\theta)$ is as defined in (5), C_i is the choice set for consumer i, and

$$d_{ict} = \begin{cases} 1 & \text{if consumer } i \text{ chooses } c \text{ at time } t \\ 0 & \text{otherwise.} \end{cases}$$

We assume that at true parameters θ^0 the following population moments hold:

$$E\left[d_{ict} - P_{ict}(\theta) | \, Z^1_{ict} \right] = 0$$

where Z^1_{ict} is a vector of instruments of dimension g^1 that vary exogenously over consumers i, choices c, and time t. These are described later in this section. We similate the choice probability using R draws to yield $P^R_{ict}(\theta)$ which, for any $R \geq 1$, by construction has the property of being an unbiased estimate. Thus we have g^1 sample moment conditions as follows:

$$\sum_i \sum_t \sum_{c \in C_{it}} \left[d_{ict} - P^R_{ict}(\theta) \right] Z^1_{ict}. \tag{6}$$

The second set of moments uses residuals based the difference between unbiased simulators of continuous choice predictions of the model and the observed data. The model predicts continuous quantities at the level of the individual store for all 30 stores in J_i and for each category k. No consumer has positive quantities at more than two stores so most of these predictions are zeros. To avoid using most of these zero observations, the continuous choice moment conditions use the top two expenditure predictions for each category: define the store at which the consumer spends the greatest total amount (across all categories) in period t as store a and the other store as b. We use the terms major store and minor store to refer to these. We denote the quantities in these stores as q^a_{ikt} and q^b_{ikt} for each category k. (For the purposes of these definitions if a consumer chooses $j = 0$—the option of no second store—then $j = 0$ is classed as b and $q^b_{ikt} = 0$ for all k.) More formally the expectations of q^a_{ikt} and q^b_{ikt} are given as follows by the model:

$$E\left(q_{ikt}^{a},\theta\right)=\sum_{c}\sum_{j,j'\in c}\int\tilde{q}_{ijckt}(\nu_{ickt})1_{\left[r_{ijct}>r_{ij'ct}\right]}\Pr(w_{ict}(v_{it})>w_{ic't}(v_{it})\ \forall c|\,\nu_{it})dF\left(\nu_{it}\right). \tag{7}$$

$$E\left(q_{ikt}^{b},\theta\right)=\sum_{c}\sum_{j,j'\in c}\int\tilde{q}_{ijckt}(\nu_{ickt})1_{\left[r_{ijct}<r_{ij'ct}\right]}\Pr(w_{ict}(v_{it})>w_{ic't}(v_{it})\ \forall c|\,\nu_{it})dF\left(\nu_{it}\right). \tag{8}$$

where r_{ijct} is the consumer's total spending in store j:

$$r_{ijct}=\sum_{k=1,\ldots,6}\left(p_{jkt}^{d(i)}q_{jk}\right)$$

and $1_{\left[r_{ijct}>r_{ij'ct}\right]}$ is an indicator that equals 1 if the inequality is true and 0 otherwise (if there is a tie the stores are classified arbitrarily).

We use moments that match the model's predictions for $\left(q_{itk}^{a},q_{itk}^{b}\right)$ as defined above with the observed data. These moments are informative about the store and consumer attributes that influence total expenditure in each category, and how it is split between the consumer's two stores. We assume that at true parameters θ^{0} the following population moments hold for each k:

$$\begin{aligned}
E\left[q_{ikt}^{a}-E\left(q_{ikt}^{a},\theta\right)\middle|\,Z_{it}^{2,a}\right] &= 0\\
E\left[q_{ikt}^{b}-E\left(q_{ikt}^{b},\theta\right)\middle|\,Z_{it}^{2,b}\right] &= 0\\
E\left[\left(q_{ikt}^{a}\right)^{2}-E\left(\left(q_{ikt}^{a}\right)^{2},\theta\right)\middle|\,Z_{it}^{2,a}\right] &= 0\\
E\left[\left(q_{ikt}^{b}\right)^{2}-E\left(\left(q_{ikt}^{b}\right)^{2},\theta\right)\middle|\,Z_{it}^{2,b}\right] &= 0\\
E\left[1\left[q_{ikt}^{a}>0\right]-E\left(1\left[\tilde{q}_{ikt}^{a}(\nu_{ickt})>0\right],\theta^{0}\right)\middle|\,Z_{it}^{2,a}\right] &= 0\\
E\left[1\left[q_{ikt}^{b}>0\right]-E\left(1\left[\tilde{q}_{ikt}^{b}(\nu_{ickt})>0\right],\theta^{0}\right)\middle|\,Z_{it}^{2,b}\right] &= 0.
\end{aligned}$$

where $Z_{it}^{2,a}$ and $Z_{it}^{2,b}$ is a vector containing g^{2} instruments that vary exogenously across i and t. We discuss the instruments in more detail below. The simulated sample moments that are minimized substitute $E^{R}\left(\left(\tilde{q}_{ikt}^{a}(\nu_{ickt})\right)\right.$ etc., in the above. Togetheer these six moment conditions yield another $6 * g^2$ moments for each k.

Finally, we include a moment that is informative about which store attributes result in a store being the major store out of any pair c. We generate a prediction for the size (sales area) and distance (from consumer) of the 'major' and 'minor' store in the chosen pair.

$$
\begin{aligned}
E\left(\text{area}_{it}^{a},\theta\right) &= \sum_{c}\sum_{j,j'\in c}\int \text{area}_{j}.1_{\left[r_{ijct}>r_{ij'ct}\right]}\Pr(w_{ict}(v_{it})>w_{ic't}(v_{it})\ \forall c|\,\nu_{it})dF\left(\nu_{it}\right)\\
E\left(\text{area}_{it}^{b},\theta\right) &= \sum_{c}\sum_{j,j'\in c}\int \text{area}_{j}.1_{\left[r_{ijct}<r_{ij'ct}\right]}\Pr(w_{ict}(v_{it})>w_{ic't}(v_{it})\ \forall c|\,\nu_{it})dF\left(\nu_{it}\right)\\
E\left(\text{dist}_{it}^{a},\theta\right) &= \sum_{c}\sum_{j,j'\in c}\int \text{dist}_{ij}.1_{\left[r_{ijct}>r_{ij'ct}\right]}\Pr(w_{ict}(v_{it})>w_{ic't}(v_{it})\ \forall c|\,\nu_{it})dF\left(\nu_{it}\right)\\
E\left(\text{dist}_{it}^{b},\theta\right) &= \sum_{c}\sum_{j,j'\in c}\int \text{dist}_{ij}.1_{\left[r_{ijct}<r_{ij'ct}\right]}\Pr(w_{ict}(v_{it})>w_{ic't}(v_{it})\ \forall c|\,\nu_{it})dF\left(\nu_{it}\right).
\end{aligned}
$$

The following four population moments are assumed to hold at the true θ:

$$
\begin{aligned}
E\left[\text{area}_{it}^{a}-E\left(\text{area}_{it}^{a},\theta\right)\right] &= 0 \qquad E\left[\text{area}_{it}^{b}-E\left(\text{area}_{it}^{b},\theta\right)\right]=0\\
E\left[\text{dist}_{ikt}^{a}-E\left(\text{dist}_{it}^{a},\theta\right)\right] &= 0 \qquad E\left[\text{dist}_{it}^{b}-E\left(\text{dist}_{it}^{b},\theta\right)\right]=0
\end{aligned}
$$

We simulate the expected size and distances, denoted $E^{R}\left(\text{size}_{it}^{a},\theta\right)$ etc., setting size and distance to zero when the consumer chooses $j=0$. The sample analogues of these moments are

$$
\begin{aligned}
&\sum_{i}\sum_{t}\left[\text{area}_{it}^{a}-E^{R}\left(\text{area}_{it}^{a},\theta\right)\right], \qquad \sum_{i}\sum_{t}\left[\text{area}_{it}^{b}-E^{R}\left(\text{area}_{it}^{b},\theta\right)\right]\\
&\sum_{i}\sum_{t}\left[\text{dist}_{it}^{a}-E^{R}\left(\text{dist}_{it}^{a},\theta\right)\right], \qquad \sum_{i}\sum_{t}\left[\text{dist}_{it}^{b}-E^{R}\left(\text{dist}_{it}^{b},\theta\right)\right].
\end{aligned}
$$

All the simulated expectations, denoted $E^{R}\left(q_{ikt}^{a},\theta\right)$ etc, use R draws from consumer taste shocks. For any $R\geq 1$ these simulators are unbiased by construction as we draw randomly from the full distribution of consumer tastes and do not condition on any subset of consumers. To minimize computational costs we use one draw per consumer. As there are many consumers we end up taking thousands of draws from the support of $F(v)$.

The moments above are combined to give a population moment condition $m(\xi, \lambda) = 0$. As $m()$ is nonlinear in parameters (ξ, λ)—and the firm-category-demographic dummies ξ_{fk}^{d} (for $f = 1, ..., 12$; $k = 1, .., 6$; and $d = 1, .., 8$) have a high dimension—it is useful to find ways to avoid estimating the full set of parameters using them. We contract the dimension of the nonlinear search to the dimension of λ and estimate the ξ using a set of moments that are linear. This is done using an extra exactly-identified moment condition that is linear in the parameters that provides a unique estimate $\hat{\xi}^t = \xi(\lambda^t)$ for any candidate value λ^t. These estimators can be substituted into the nonlinear moment conditions in advance for any candidate value of λ, i.e. $m(\xi(\lambda^t), \lambda^t) = m(\lambda^t)$. The third moment providing the estimator $\xi(\lambda^t)$ uses the residuals ν_{ijkt} from the category demand functions, conditional on positive category demands and on choice of store. Conditioning on $q_{jkt} > 0$ ensures that the demands are derived from the first-order condition for an interior optimum (3), which provides an expression for the residuals that is linear in ξ:

$$
\begin{aligned}
\nu_{ijkt}(\xi|\lambda^t, q, x) &= (\gamma_k^t q_{jkt} + \sigma_{ck}^t q_{j'kt} + \alpha_i^t p_{jkt} - \gamma_x^t x_{jk}) - \sum_{f,d} \xi_{f(j)k}^{d(i)} 1_{[f(j)=f]} && (9) \\
&= g(q_{jkt}, q_{j;kt}, p_{jkt}, x_{jk}, \lambda^t) - \sum_{f,d} \xi_{f(j)k}^{d(i)} 1_{[f(j)=f]}. && (10)
\end{aligned}
$$

where $1_{[f(j)=j]}$ is an indicator taking the value 1 if $f(j) = f$ and zero otherwise.

For any candidate λ^t the equation (9) can be interpreted as a regression of $g()$ on regressors $1_{[f(j)=f]}$ with residual the ν_{ijkt}. However, as condition on positive expenditure, the expected value of the residual above is not independent of the regressors . To allow for this dependence we use a set of instruments Z^{3k} that are correlated with the $1_{[f(j)=f]}$ but not with v_{ijkt}. The instruments are described below and exactly identify the parameters $\xi^{d(i)}_{f(j)k}$. The moment conditions are

$$\sum_i \sum_t \sum_{\{j: j\in c \text{ and } q_{ijkt}>0\}} \left[g(x,q,\lambda^t) - \sum\nolimits_{f,d} \xi^{d(i)}_{f(j)k} 1_{[f(j)=f]} \right] Z^{3k}_{it} = 0 \text{ for each } k. \qquad (11)$$

The parameters $\xi^{d(i)}_{f(j)k}$ are normalised so that they have mean zero before being used in the moments $m()$. [5] For Z^{3k}_{it} we use instruments based on choice set variation across consumers. The probability of demographic type d making a positive purchase of category k in store j (and j' respectively) can be estimated as a function of exogenous variables X_{it} in a first step in a reduced form model to give probabilities $\hat{P}^{d(i)}_{ijk}(X_{it})$, where X_{it} are uncorrelated with v_{ijt} but vary across consumers (depending on local choice sets) and therefore are correlated with $1_{[f(j)=f]}$. This type of instrument is used in the discrete-continuous model in Dubin and McFadden (1984). The ξ's are exactly identified by (11) as there is one $\hat{P}^{d(i)}_{ijk}(X_{it})$ probability for each firm f, category k , and demographic type $d(i)$.

The data do not in all cases give an individual store choice. For some firms the firm is known but not the store. In these cases case we aggregate the moments to be matched to the level of the firm, and replace j with f subscripts. So we use predictons for the firms that are chosen in the first set of moments (instead of the stores that are chosen) and in the moments for the expenditure (by category) we match expenditure to the firm (rather than store). In the moments matching size and distance we use the *nearest* of the alternative possible stores (of the same firm) in the observed data and match this to the nearest store of the firm of the store that is predicted by the model. In the conditional moments for ξ where we don't observe store, which is the case for a number of fascia, we can either (i) drop observations where there is significant ambiguity about x_{jk} —i.e. consumers who have stores of significantly different size from the same fascia for the fascia whose stores we do not observe in the demand data—or (ii) use a weighted estimate of x_{jk} based on reduced form probabilities estimated in a

[5] This adjusts for the fact that the assumption $E(v|z) = 0$ does not imply $E(v|z) = 0$. We specify x_{jk} in (9) to contain a constant for each category.

previous stage. [A related point: in some of these cases we do know that two distinct stores of the firm were visited without knowing how to match these stores to those in the choice set. But this information appears relatively infrequently and we do not use it as it would add complication].

Before discussing instruments $\{Z^1_{ict}, Z^{2,a}_{it}, Z^{2,b}_{it}\}$ individually, we note that the main identification issue in choosing instruments arises from the potential endogeneity of price. Recall that our observed prices are national so we do not have an issue regarding market-specific demand shocks. As our model includes firm dummies ξ for each category, we have controlled for unobserved utility that varies across firms (cross-sectional unobserved utility). However these effects are fixed over time so that we have not controlled for any time-specific demand shocks. Thus observed price for any firm may be correlated with unobserved time-specific shocks. To instrument for variation over time in the prices we use instruments for price based on data that are correlated with marginal costs but assumed not to be correlated with demand shocks. Specifically, these instruments are: the Euro-£ exchange rate, which affects the prices of certain acricultural products; energy prices which affect the price of some of the categories (especially chilled and frozen foods) more than others; agricultural commodity prices in other EU countries (e.g. milk, fruit, meat, and wheat); and retail prices in other EU countries for equivalent categories. (The last of these uses the assumption in Nevo (2001) and Hausman (1998) that prices in different markets for the same product are correlated with common cost shocks but not with demands shocks).

Recall that the price parameters in the model—those that govern the effect of a price change on the demand for each category at any store—enter at two levels. First, at the level of store choice, parameter α_i governs the effect of the overall price index on store choice. Second, the effect of price on demand for each category k, conditional on for store-level choice, is determined by the parameters γ_k (for $k = 1, .., 6$). With respect to the upper level we note that identification of α is helped by the presence of the following sources of exogenous variation in overall prices across consumers at any given point in time: (i) consumers in different locations have different stores (with different firms and therefore prices) in their choice sets, and (ii) consumers of different demographic types attach different weights to each category k, resulting in variation across consumer types in the overall price for any firm. This latter point means that consumer type dummies serve as instruments for the overall price. With respect to the lower level we note that the presence of category-specific firm dummies means that identification of γ_k relies upon variation over time in the cost instruments outlined in the previous paragraph.

We now discuss in detail the instruments used for each set of moment conditions. We have already discussed the instruments Z_{it}^{3k} so we discuss the other instruments namely $\{Z_{ict}^1, Z_{it}^{2,a}, Z_{it}^{2,b}\}$. We begin with the instruments $Z_{itk}^{2,a}$ for predictions based on continuous category demands at the consumer's major store. These instruments are (i) consumer demographics used anywhere in the utility model including consumer type dummies; (ii) the expected values for all the non-price attributes of the consumer's (major store) including firm dummies based on a reduced form choice probability model for the a store; and (iii) category-specific instruments for price based on the cost side instruments detailed above. Equivalently for $Z_{itk}^{2,b}$.

Finally the Z_{ict}^1 instruments (used in the predictions for choice c) include: all observable non-price store variables for stores (j, j') in c used anywhere in the utility function including firm dummies; all consumer demographics used anywhere in the utility function including a dummy for the consumer's demographic type; and cost-side instruments for the prices of each of all k categories as outlined in the previous paragraph.

5 Results and Counterfactual Analysis

The parameters are presented in Table 7. The parameters are largely of the expected sign, with the possible exception of *quadrant* which we had expected to be positive (suggesting that people prefer to combine stores in the same quadrant). The parameter on number of product lines shows that the demand increases with the number of product lines. The sign on σ should be interpreted as indicating that two stores with a close product overlap are seen as closer substitutes, as we would expect. The parameters on distance and store size are of the expected signs.

The remaining tables illustrate the elasticities and counterfactual analysis. Table 8 shows that the average price elasticities are about -1 for one stop shoppers (conditional on choice of store) and about -4 for two stop shoppers. This suggests that the presence of two stop shoppers has a strong downward effect on market power. Table 9 shows the price elasticities of discrete choice –the proportionate effect of a change in row price on the number of people choosing to visit column firm stores. The Table 10 shows the price elasticities of overall demand combining discrete and continuous choices. It is especially interesting to note that those stores that have a high proportion of two stop shoppers —e.g. M&S and the Discounters—have the biggest difference when these elasticities are compared with the discrete-only elasticities, showing that avoiding two stop shoppers is a route to market power. Table 11 shows the effect of adding a hypothetical big-box Tesco (size 50,000 sq ft) in Oxford. The model is scaled up to the population level using census data. The effects on rival stores is shown. It can be seen that the Tesco takes demand not just from other firms (such as Sainsbury) that specialize in big box stores but also on the smaller stores of firms that locate on High Streets, such as Somerfield , Waitrose , Iceland, and Coop.

6 Conclusions

The paper specifies and estimates a model of supermarket demand that allows us to analyze important policy issues relating to planning and store format.

7 Appendix on Price Index Construction

A price index $P_{jt}^{d(i)}$ is constructed for each week $t = 1, .., 156$, consumer group $d = 1, ..., 8$, and chain $f = 1, ..., 17$ in the choice set J. P_{jt}^{d} price index is designed to allow comparisons of the price of buying a basket of goods *both* (i) over weeks for a given chain and (ii) across chains for a given week. The intertemporal aspect is important when estimating the effect on continuous demand of price changes over time. The inter-chain aspect is important in evaluating the effect of cross sectional choice set variation on observed choice of chain, and the role of price differences on that choice.

Chain prices, rather than individual store prices, can be used because of the practice of national pricing which is typical in the main firms. Chains are here distinguished by the firm operating the store and, where firms operate heterogeneous-sized stores, firm and size of store (i.e. we treat Tesco Medium as one chain, Tesco Large a separate chain, etc). We compute a price for each store size band because pricing variation occurs across store sizes for given firms—e.g. Tesco Metro (a small size format) has a different price list from standard-sized Tescos (these price lists are applied nationally).

Consumer groups d are classified by social class groups and household size classes and are used to allow price indices to be constructed relevant to the purchase patterns of each demographic group.

To facilitate comparisons over time conditional on chain j, we use a consistent set of products for each f through the 156 weeks, and we use weights that are not time-varying. As the weights are not weighted using first period quantities but are instead weighted using expenditure averages over the entire 3 years we do not have the upward "substitution biases" associated with traditional Laspeyres indices.

In making comparisions across chains we exploit the fact that they all sell the same product *categories* with only minor exceptions: there are 183 such categories (representing 96% of spending) that are sold in all 17 choices. These categories are defined quite narrowly: fruit, tea, breakfast cereal, eggs, etc.[Note that these categories are much narrower than the 6 broad categories used in the paper and denoted k]. If we are given a vector of *category prices* p^d_{aft} for each chain f then it is straightfoward to aggregate them to the scalar P^d_{ft}: we can weigh them using weights determined by national expenditure proportions on each category a. The use of national rather than store specific weights allows comparisons that are based on price differences, rather than differences in weights, across choices. It also reduces "substitution bias" as store specific weights would over-weigh the products that get relatively low prices in each chain. A further problem in the store choice context is the *selection bias* problem, namely that store specific weights weigh products according to the tastes of those who actually choose the store, and these may not be the correct weights for people who do not do this. Thus the avoidance of store specific category weights on category prices avoids this problem.

The main challenge is in the construction of the *category prices*—e.g. the price of 1kg of breakfast cereal or 1L of milk at chain f—because a substantial number of products within each category are not common across stores. In the example of breakfast cereals, private label products are unique to the firm selling them while national brands such as Kelloggs Corn Flakes are common across almost all stores. The problem here is *not* that the products are of different quality (i.e. that some own-label products may be of higher quality at one chain than another). The problem is that for many products there is no exact counterpart in the alternative stores so we cannot form a "matched" set of common products and apply a common weight based on national revenues for that product. This problem is a limited one for the following three reasons. First, the categories are quite narrow (e.g. eggs, instant coffee, and liquid milk are examples of categories) so the within-category weighting is relatively less important than the inter-category weights (for which as noted in the previous paragraph we *do* have direct comparisons and national revenues). Second we weigh products within categories using national revenues where possible, namely on national brands, and use store-specific revenue data only where there is no alternative, namely on the private label products. Third, we construct separate indices for each of 8 demographic groups of consumers, which should reduce the taste heterogeneity of the consumers within any group, and thus the extent of the underlying selection bias problem.

The raw data consist of product prices p_{hft} and volumes q_{hft} based on records of consumer purchases over time; prices are expressed per unit of volume. We observe thousands of products $h = 1, ...,$ in product categories $a = 1, ..., A$ where $A = 183$. Within each category, products are measured in comparable quantity units (e.g. litres, kg, etc) so quantity q_{hft} can be meaningfully added for all h in any a. Distinct categories on the other hand are not in comparable quantity units so quantity aggregation is not possible across categories, so instead we use revenue weights at that level. The price index is thus done in two steps as follows.

1. **Construction of category prices from product prices.** We construct for each chain f a category price index p^d_{atf} using the prices of individual products h in category a as follows

$$p^d_{atf} = \sum_{h \in a} w^d_{hf} p_{hft}$$

where the weights w^d_{hf} capture the importance of product h in category a. The weights are common over time but they vary over chains, as not all products are stocked in all chains (as we noted above). Each firm has some products only stocked at that chain—namely private label products. The extent of private label products varies from firm to firm. To get w^d_{hf} we first construct product h's share of the volume of category a sold in the set of chains that that stock product h:

$$\kappa^d_{hf} = \frac{\sum_{f=1}^{17} 1_f[h] \sum_{t=1}^{156} q^d_{ht}}{\sum_{h \in a} \sum_{f=1}^{17} 1_f[h] \sum_{t=1}^{156} q^d_{ht}}$$

for each h in a, where $1_f[h]$ is an indicator for whether chain f stocks product h. Thus in the case of products that are stocked in all chains, κ^d_{hf} is product h's share of total category a volume whereas on the other hand in the case of private label products κ^d_{hf} is h's share of the volume that chain f sells of category a, with analogous weights for intermediate cases of products stocked in more than one chain but less than all chains. We then obtain w^d_{hf} by scaling κ^d_{hf} to ensure that the weights add up to unity for each chain f as follows:

$$w^d_{hf} = \frac{\kappa^d_{hf}}{\sum_{h \in a} \kappa^d_{hf}}.$$

This w^d_{hf} weighs products in proportion to κ^d_{hf}. Note that this method weighs products using information that is not specific to firm f were possible, which has the advantage of avoiding the substitution and selection bias issues mentioned above. For examples of how the weights work consider the following three examples:

(a) The category a price for a chain f that stocks only products that are available in all chains is constructed using weights that are in proportion to the national sales of those products

(b) The category a price for a chain that stocks only private products is constructed using weights that in proportion to those products' volumes at firm f

(c) A category a price of a chain f with a mix of national and private products—the typical case—is constructed using weights that are proportional to their individual κ's as defined above. This can be illustrated with the following example: consider a category a with three products at chain f. Suppose product 1 is a private good comprising 30% of volume of that chain's category a. Product 2 is a national brand with 50% of category volume for stores stocking it; and product 3 is also national with a 30% share of category volume for stores stocking it. The weights w^d_{hf} allocated to each product under the above scheme are 27.2%, 45.4% and 27.2% respectively. The corresponding category price p^d_{atf} that is generated by this weighting scheme falls between the two that which would be obtained in the following two alternative schemes: (a) Set private label product weights equal to firm-specific expenditure shares and set remaining products' weights in proportion to national expenditure shares (remaining products' weights being scaled to ensure all the chain's weights aggregate to unity); in the example this would yield weights of 30%, 44%, 26% i.e. weighting the private label more than our proposed scheme above. (b) Set national product weights equal to national expenditure share and set remaining private label weights in proportion to firm-specific shares (scaled to ensure weights add to unity); in the above example this would yield 20%, 50%, 30% i.e. weighing the national products more than our proposed scheme.

2. **Construct chain price index (for each of the six large categories k) from from category a prices.** Using the category prices p^d_{atf} obtained in stage 1 we construct the overall price index P^d_{ft}. This is a revenue-weighted sum of price relatives where prices are relative to p^d_{a11} i.e. price at $t = 1$ in $f = 1$ (Large ASDA). The use of price relatives ensures that the index is independent of the units chosen in any category—i.e. changing milk from pints to liters will not change the index.
$$P^d_{ft} = \sum_{a=1}^{A} \omega^d_a \frac{p^d_{aft}}{p^d_{a11}}$$
The weighting term ω^d_a is for consumer defined as follows
$$\begin{aligned} \omega^d_a &= \frac{\sum_{f=1}^{17} \sum_{t=1}^{156} p^d_{atf} q^d_{atf}}{\sum_{a=1}^{A} \sum_{f=1}^{17} \sum_{t=1}^{156} \left[p^d_{atf} q^d_{atf} \right]} \\ &= \frac{\text{national spending on category } a \text{ by group } d}{\text{total national spending by group } d} \end{aligned}$$
These weights sum to 1 for each d, i.e. $\sum_a \omega^d_a = 1$. Note that for any d the weights are constant over time t and across choices f and that the set A consists of all categories that are sold in all choice set options $f \in J$. This is done separately for all the a in each $k = 1, .., 6$.

Further practical issues:

- Where two chains f are from the same firm (e.g. Tesco Medium and Tesco

Large) we have used a common set of products and weights for each chain so that differences between the two price indices are entirely a result of differences in prices, not differences in weights or differences in included products.

- The price series p_{hft} for some h and f are incomplete over time t. That is, some products are not observed every period. To facilitate intertemporal comparisons we wish to use the same set of products each period for each a and f. To deal with the problem of missing prices in certain periods we drop products that appear in less than 40 of the 156 weeks and in less than 4 of the 4 years covered in the data and, for products that remain, we *impute* remaining missing prices using prices for chain f and product h observed in the same quarter and chain (or, failing that in a few cases, the same half year).

- The eight demographic groups d are based on four social class categories and two household size categories. Social classes correspond to occupational group and therefore to household income. Social classes 1 & 2 are small and therefore grouped together into a single group, classes 3 & 4 are each large and therefore treated as two separate groups, and 5 & 6 are small and aggregated into a fourth group. For each of these 4 social class groups we have two household size groups: small households (1 or 2 people) and large households (3 or more people).

References

[1] Dubin J A and D L McFadden (1984) "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption" *Econometrica*

[2] Gentzkow (2007) "Valuing New Goods in a Model with Complementarity" *American Economic Review*

[3] Holmes Thomas (2011) The Diffusion of Wal-Mart and Economies of Density *Econometrica* Volume 79, Issue 1, pages 253–302.

[4] Jia, P. (2008) “What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry,” *Econometrica*, 76, 1263–1316.

[5] McFadden D L (1989) “A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration” *Econometrica* 57, 995-1026

[6] Smith, Howard (2004) Supermarket Choice and Supermarket Competition *Review of Economic Studies*

Fascia	Store Size	Stores	Store Size	Market Share	Market Share	Spending	Range	price
	Class	#	Avg. (Sq. Ft)	Trips	Expenditure	Per Customer	#Lines	
ASDA		263	45411	12.49	18.10	25.28		
	L	82	32020	3.08	4.23	23.96	34405	1.01
	XL	181	51477	9.41	13.87	25.71	39794	1.01
Morrisons		294	30661	6.32	8.65	23.93		
	L	261	29038	5.06	6.81	23.51	36014	1.04
	XL	33	43498	1.26	1.84	25.60	28608	1.03
Sainsbury's		502	29431	11.65	15.44	23.16		
	M	106	6999	1.00	0.75	13.15	24405	1.18
	L	145	22985	3.51	3.78	18.82	36470	1.19
	XL	251	42628	7.14	10.91	26.69	42574	1.20
Tesco		975	23579	21.28	28.58	23.44		
	M	446	4391	4.02	3.57	15.52	38078	1.11
	L	310	26742	8.64	11.51	23.25	42759	1.12
	XL	219	58180	8.62	13.50	27.34	44956	1.12
Discounter		484	7842	6.56	4.52	12.03	18183	0.82
Iceland		621	4863	3.90	2.20	9.83	11560	1.17
Co-op		1599	4247	7.72	3.49	7.90	24512	1.26
Somerfield		793	8608	5.35	3.33	10.88	31680	1.22
Other		886	9813	19.26	11.69	10.59	30453	1.12
Marks & Spencer		284	8655	3.35	1.96	10.20	9749	1.92
Waitrose		165	19203	2.14	2.03	16.56	23493	1.48

Table 1: **Descriptive Statistics: Store & Shopping Characteristics.** For each type of store the number of stores, average store size, market shares and average expenditure is reported. For the four biggest supermarket chains the stores are split into different size bands (see text for more detail). Market shares are calculated based on the total number of trips and alternatively on the consumers' overall expenditure. Market shares and average expenditure are calculated over the sample period 2002-2005.

	ASDA		Co-op	Disc.	Iceland	M&S	Morrisons		Other	Sainsbury			Som.	Tesco			Wait.
	L	XL					L	XL		L	M	XL		L	M	XL	
ASDA L	1.00	0.96	0.34	0.20	0.17	0.03	0.42	0.37	0.42	0.42	0.30	0.46	0.39	0.46	0.43	0.48	0.29
ASDA XL	0.83	1.00	0.32	0.19	0.16	0.02	0.40	0.34	0.40	0.40	0.28	0.45	0.37	0.46	0.42	0.47	0.27
Co-op	0.47	0.51	1.00	0.26	0.23	0.04	0.50	0.44	0.45	0.48	0.36	0.51	0.50	0.53	0.50	0.54	0.35
Discounter	0.38	0.42	0.34	1.00	0.22	0.04	0.41	0.35	0.38	0.36	0.27	0.39	0.39	0.40	0.37	0.41	0.26
Iceland	0.50	0.54	0.49	0.35	1.00	0.04	0.53	0.48	0.49	0.48	0.38	0.52	0.54	0.53	0.51	0.53	0.35
Marks & Spencer	0.09	0.10	0.10	0.08	0.05	1.00	0.10	0.08	0.15	0.09	0.07	0.10	0.10	0.10	0.10	0.11	0.09
Morrisons L	0.40	0.44	0.34	0.21	0.17	0.03	1.00	0.76	0.41	0.40	0.29	0.44	0.43	0.45	0.42	0.47	0.29
Morrisons XL	0.44	0.48	0.38	0.22	0.19	0.03	0.96	1.00	0.45	0.44	0.33	0.47	0.46	0.48	0.46	0.49	0.32
Other	0.47	0.52	0.36	0.23	0.18	0.05	0.48	0.42	1.00	0.52	0.38	0.57	0.44	0.59	0.55	0.60	0.33
Sainsbury L	0.40	0.44	0.32	0.18	0.15	0.02	0.40	0.35	0.43	1.00	0.64	0.96	0.38	0.48	0.44	0.49	0.31
Sainsbury M	0.42	0.45	0.36	0.20	0.18	0.03	0.42	0.39	0.47	0.96	1.00	0.98	0.41	0.48	0.46	0.49	0.34
Sainsbury XL	0.37	0.42	0.30	0.17	0.14	0.02	0.37	0.32	0.41	0.82	0.56	1.00	0.35	0.46	0.42	0.47	0.29
Somerfield	0.43	0.47	0.39	0.22	0.20	0.03	0.49	0.41	0.42	0.43	0.32	0.47	1.00	0.54	0.50	0.55	0.31
Tesco L	0.37	0.43	0.30	0.17	0.14	0.02	0.38	0.32	0.42	0.41	0.28	0.46	0.40	1.00	0.83	0.93	0.28
Tesco M	0.39	0.44	0.32	0.18	0.15	0.03	0.40	0.35	0.44	0.42	0.30	0.47	0.42	0.94	1.00	0.95	0.29
Tesco XL	0.36	0.41	0.29	0.16	0.14	0.02	0.38	0.31	0.41	0.39	0.27	0.45	0.39	0.88	0.81	1.00	0.27
Waitrose	0.42	0.46	0.36	0.20	0.17	0.04	0.44	0.39	0.42	0.48	0.36	0.52	0.42	0.51	0.48	0.52	1.00

Table 2: **Product Overlap.** Each Cell shows proportion of the products in row store that are available in a column store.

	Share (%)	Cumul. (%)
Households in Each Shopping Type		
One Store per Week	61	61
Two Stores per Week	26	87
One Trip per Week	53	53
Two Trips per Week	26	79
Share of All Spending		
Largest Weekly Store	86	86
Second Largest Weekly Store	11	97
Largest Weekly Trip	79	79
Second Largest Weekly Trip	15	93
Two-Trip Shoppers only:		
Proportion that go to Same Firm	35	
Proportion that go to Same Store	32	

Table 3: **Descriptive Statistics: One-Stop and Two-stop shopping.** Calculated over the sample period 2002-2005.

	Big 4				Lower Quality		Mid Range			High Quality		
	ASDA	Morrisons	Sainsbury's	Tesco	Discounter	Iceland	Co-op	Somerfield	Other	Marks & Spencer	Waitrose	Only One Trip
ASDA	0.05	0.83	0.60	0.58	1.08	0.86	0.77	0.59	0.92	0.85	0.38	1.83
Morrisons	0.83	0.04	0.45	0.53	1.32	0.65	0.80	0.49	1.00	0.94	0.32	1.72
Sainsbury's	0.60	0.45	0.15	0.73	0.85	1.11	0.77	0.86	0.88	1.88	1.38	1.62
Tesco	0.58	0.53	0.73	0.14	0.95	1.01	0.87	0.88	0.88	1.14	0.93	1.81
Discounter	1.08	1.32	0.85	0.95	0.00	1.67	1.21	1.22	1.55	0.63	0.71	0.84
Iceland	0.86	0.65	1.11	1.01	1.67	0.00	1.20	1.72	1.69	1.49	1.36	0.57
Co-op	0.77	0.80	0.77	0.87	1.21	1.20	0.00	1.42	1.52	1.08	0.89	1.08
Somerfield	0.59	0.49	0.86	0.88	1.22	1.72	1.42	0.00	1.23	1.02	1.00	1.15
Other	0.92	1.00	0.88	0.88	1.55	1.69	1.52	1.23	0.00	1.33	0.82	1.24
Marks & Spencer	0.85	0.94	1.88	1.14	0.63	1.49	1.08	1.02	1.33	0.00	2.72	0.59
Waitrose	0.38	0.32	1.38	0.93	0.71	1.36	0.89	1.00	0.82	2.72	0.00	1.23
Only One Trip	1.83	1.72	1.62	1.81	0.84	0.57	1.08	1.15	1.24	0.59	1.23	0.00

Table 4: **Descriptive Statistics Two-Stop Shopping: Supermarket Chains.** The table reports the likelihood of a particular fascia combination being chosen relative to the probabilities of each choice: $\frac{P(AB)}{P(A)*P(B)}$. If the two visits are independent events, then $P(AB) = P(A) * P(B)$ and the value in the table for this combination will be equal to one. In case that the stores are likely (unlikely) to be combined a value greater (smaller) than one is reported. Likely combinations are highlighted with bold font. Note that we aggregate trips at the level at which we model the options in the choice set. For the big four we treat differently sized stores as different options and combinations of stores within the same fascia are therefore possible. For all other fascias the diagonal elements are zero by construction.

	< 15,000 sq feet	< 40,000 sq feet &		Only One
		> 15,000 sq feet	> 40,000 sq feet	Trip
< 15,000 sq feet	0.76	1.09	0.99	1.20
< 40,000 sq feet &	1.09	0.25	0.51	1.62
> 15,000 sq feet				
> 40,000 sq feet	0.99	0.51	0.23	1.76
Only One Trip	1.20	1.62	1.76	0.00

Table 5: **Descriptive Statistics Two-Stop Shopping: Size Bands.** The table reports the likelihood of a particular size combination being chosen relative to the probabilities of each choice: $\frac{P(AB)}{P(A)*P(B)}$. If the two visits are independent events, then $P(AB) = P(A) * P(B)$ and the value in the table for this combination will be equal to one. In case that the stores are likely (unlikely) to be combined a value greater (smaller) than one is reported. Likely combinations are highlighted with bold font. Note that we aggregate trips at the level at which we model the options in the choice set. Combinations of stores within the same size band are therefore possible.

Category	DISCOUNTERS	M&S	ICELAND	BIG 4	WAITROSE	NEIGHBORHOOD
Alcohol	11.5%	3.8%	3.4%	7.8%	6.3%	10.1%
Chilled	13.9%	51.2%	11.6%	19.7%	23.3%	18.4%
Dry	28.5%	12.0%	17.2%	24.1%	22.9%	25.2%
Fresh	26.4%	28.0%	15.0%	27.4%	31.7%	29.0%
Frozen	8.2%	2.7%	50.0%	6.9%	4.9%	7.1%
Household	11.5%	2.4%	2.8%	14.1%	10.9%	10.1%

Table 6: **Share of Category in Format Sales.** The table reports share of each format's revenue from each category. ASDA, Morrison, Safeway, Sainsbury, Tesco are classed as Big Four. Somerfield and Coop are Neighbourhood stores. Aldi, Lidl, and Netto are classed as Discounters .

		k	Estimate	Standard Error
Second order terms	γ_{1k}	Alcohol	4.76	(1.15)
		Chilled	3.00	(1.00)
		Dry	3.06	(1.05)
		Fresh	9.00	(2.35)
		Frozen	13.43	(5.10)
		Household	5.50	(1.50)
Family size	γ_{2k}	Alcohol	1.30	(0.09)
		Chilled	6.83	(1.21)
		Dry	7.87	(0.95)
		Fresh	14.30	(3.21)
		Frozen	10.83	(2.21)
		Household	13.39	(3.45)
Managerial job	γ	Alcohol	4.78	(1.78)
		Chilled	3.16	(0.85)
		Dry	-1.41	(0.96)
		Fresh	3.92	(1.33)
		Frozen	-5.43	(1.22)
		Household	1.41	(1.14)
number of lines	γ_x	All	25.28	(9.82)
product overlap	σ	All	2.00	(0.97)
price	α	–	21.50	(8.88)
scaling term	μ_4	–	0.10	(0.02)
distance	β_1	–	-9.00	(2.34)
sales area	β_2	–	6.25	(2.01)
same quadrant dummy	β_3	–	-0.65	(0.5)
two store shopping cost	β_4	–	16.92	(6.35)
ν_{itfk}	μ_1	–	12.71	(4.41)
ν_{if}	μ_2	–	3.10	(0.87)
υ_{ik}	μ_3	–	0.20	(0.01)

Table 7: **Utility Parameters.** Standard errors in parentheses; Obs. 77055.. Demographic effects were included in the gammas (and estimated using the nonlinear moments. Firm -category fixed effects are estimated in the linear moments xsi(lambda) and are not shown .We have not included any consumer heterogeneity in the beta parameters at this stage. The price coefficient enters as price/income to allow high income people to be less sensitive to price

	IV
Average own-price elasticity for one-stop shopper	-1.36
Average own-price elasticity for two-stop shopper	-4.04
Average cross price elasticities	4.31

Table 8: Table Elasticities Conditional on Choice of Stores

	Aldi	Asda	Coop	Lidl	Morrison	MS	Sains	Somerfield	Tesco	Waitrose
Aldi	-1.87	0.36	0.21	0.04	0.24	0.03	0.27	0.06	0.44	0.02
Asda	0.06	-1.89	0.19	0.02	0.18	0.03	0.27	0.06	0.43	0.03
Coop	0.03	0.2	-1.87	0.02	0.16	0.02	0.24	0.07	0.4	0.05
Lidl	0.08	0.31	0.25	-1.9	0.2	0.03	0.29	0.08	0.48	0.06
Morrison's	0.06	0.27	0.22	0.02	-2.13	0.04	0.28	0.07	0.43	0.04
MS	0.02	0.12	0.09	0.01	0.11	-1.6	0.26	0.02	0.26	0.08
Sainsbury's	0.03	0.21	0.17	0.02	0.15	0.04	-2.05	0.06	0.49	0.08
Somerfield	0.03	0.19	0.22	0.02	0.15	0.02	0.26	-2.22	0.47	0.06
Tesco	0.03	0.2	0.18	0.02	0.14	0.03	0.31	0.07	-1.79	0.07
Waitrose	0.01	0.1	0.13	0.01	0.09	0.05	0.36	0.05	0.44	-2.17

Table 9: Table Discrete Elasticities

	Aldi	Asda	Coop	Lidl	Morrison	MS	Sains	Somerfield	Tesco	Waitrose
Aldi	-3.6	0.56	0.35	0.04	0.36	0.05	0.43	0.1	0.74	0.04
Asda	0.1	-3.37	0.27	0.04	0.26	0.05	0.4	0.09	0.64	0.04
Coop	0.08	0.34	-3.79	0.04	0.26	0.05	0.39	0.12	0.71	0.07
Lidl	0.1	0.47	0.41	-3.48	0.29	0.05	0.44	0.12	0.8	0.1
Morrison's	0.1	0.4	0.32	0.04	-3.67	0.06	0.41	0.09	0.65	0.06
MS	0.09	0.37	0.28	0.04	0.29	-4.25	0.62	0.09	0.79	0.14
Sainsbury's	0.07	0.34	0.27	0.03	0.23	0.07	-3.7	0.1	0.75	0.12
Somerfield	0.07	0.32	0.37	0.04	0.22	0.04	0.42	-3.95	0.76	0.1
Tesco	0.07	0.32	0.29	0.03	0.21	0.05	0.45	0.11	-3.33	0.1
Waitrose	0.03	0.18	0.24	0.03	0.16	0.07	0.55	0.11	0.77	-4.07

Table 10: Table Discrete and Continuous Elasticities

	Asda	Budgen	Coop	Iceland	Morrisons	M & S	Sainsbury	Somerfield	Tesco	Waitrose
Typical format:	Big Box	High Street	High Street	High Street	Big box	High Street	Big Box	High Street	Big Box	High Street
#stores	1	11	74	7	2	4	12	18	40	8
Ex Ante Demand:										
Sales/wk (£k)	227	545	2347	336	683	44	2311	926	5899	241
#Customers/wk	5655	21529	112144	22777	17984	25371	57750	37227	143189	9503
Sales/ Cons.	40.19	25.30	20.93	14.79	38.03	1.74	40.02	24.90	41.19	25.37
Sales Share	0.02	0.04	0.17	0.02	0.05	0.01	0.17	0.06	0.43	0.017
Trip Share	0.01	0.05	0.25	0.05	0.04	0.05	0.12	0.08	0.32	0.02
Change from adding large Tesco (proportionate):										
Sales/wk	-0.37	-0.07	-0.16	-0.16	-0.01	-0.04	-0.19	-0.15	0.26	-0.13
#Customers/wk	-0.30	-0.05	-0.10	-0.10	-0.01	-0.01	-0.14	-0.12	0.24	-0.10

Table 11: **Effect of adding a 50,000 sq ft Tesco store in Oxfordshire.** The table reports the revenues, number of customers, and the changes to these for each of a number of firms serving the Oxfordshire market, and the change to these when a large big-box Tesco is added. An effect is felt on the small format (as well as the firms with large format stores). Results are preliminary and subject to change. .