Search and Price Dispersion in Online Grocery Markets

Timothy J. Richards, Stephen F. Hamilton, William Allender

September 25, 2015

Abstract

Prices for similar products often differ between retail outlets, leading consumers to actively search for products that meet their needs at the lowest possible price. Prices differ among retailers, and search intensity differs among consumers because search is a costly activity and consumers differ in their costs of search. How variety and the multiproduct nature of retailing affect search costs, search intensity, and the dispersion of prices, however, is not well understood. In this paper, we use online grocery pricing data from four retailers in the UK to estimate search costs and equilibrium price dispersions. When consumers search for single products, we find that variety reduces the cost of search and induces consumers to search less, which increases the pricing power of online retailers. However, when consumers search for multiple products, search costs still fall in variety, but consumers search more intensively across stores, potentially increasing the competitiveness of online retail markets.

1 Introduction

The fact that the law of one price is usually violated in practice (Means 1939) is well documented. That seemingly identical products sell for remarkably different prices at the same point in time may be due to the coexistence of informed and uninformed consumers (Varian 1980; Burdett and Judd 1983; Carlson and McAfee 1983), consumers with high and low valuations (Jeuland and Narasimhan 1985; Pesendorfer 2002), those that vary in store loyalty (Villas-Boas 1995), and consumers who search intensively or do not search at all (Wilde and Schwartz 1979). If a lack of information is the primary source of price dispersion, then price variation due to costly search should be lower online, and the law of one price restored. However, Brynjolfsson and Smith (2000), Clay, Krishnan, and Wolff (2001), and Goolsbee (2001) document significant price variation even in online markets that are more near to the textbook description of perfection. By removing all the physical elements of search cost (driving, parking, turning pages in a newspaper ad, for example) the remaining costs of search in online markets reflect pure mental "processing costs" that are at the core of any discussion of how consumers perceive and react to price variation. Most empirical studies of online price dispersion consider markets for single products in which the search process is arguably simplified, whereas many retail products (food, for example) are purchased in more complex, multicategory, multiproduct environments. In these environments, consumers not only make many choices within a short period of time, but typically face a wide assortment of products from which to choose. Deeper assortments, in turn, can either complicate or faciliate the search process. In this study, we examine the question of how search costs vary with the number of products available in a differentiated, multiproduct, online retail environment.

We examine this question in the context of a rapidly-growing online retail market: Online grocery purchases. The magnitude of search frictions in online markets is particularly important in the food retailing industry given the small margins earned by existing retailers, and the expected growth on online food retailing. Indeed, online markets for retail food products are currently small, but are growing rapidly. In 2010, the Hartman Group (Hartman 2012) found that online grocery sales represented only $13.0 billion of the over $1.0 trillion retail food business; however, comScore (2011) reports the online component of grocery sales is growing much more rapidly than traditional grocery sales at more than 10% per year.[1] Moreover, with Amazon and Walmart beginning to compete in the "order and deliver" market, many of the obstacles to growth (access, delivery, and price) will be removed within the next few years.[2] Because online retailing differs from traditional retailing both in terms of the cost of search and the potential to offer much greater variety online, understanding how search costs vary with variety is critical to understanding pricing decisions by online food retailers.

There is a fundamental, yet unresolved, debate regarding how search costs change with variety, or assortment depth. Because variety is a key strategic variable in retail environments (Richards and Hamilton 2014), retailers may be able to manipulate consumers' cost of search to their own advantage. If the cost of search is fixed (for example, the cost of driving to the store and finding a parking spot) then average search costs fall with the number of products considered, and deeper markets should be more efficient. Further, if greater variety implies that products are more densely packed into a given attribute space, then this finer differentiation among products makes finding the desired product simpler. Greater variety implies lower search costs within each store as less effort is required to find the desired item,

[1] Traditional retail outlets account for roughly $500.0 billion in sales, and non-conventional outlets, such as club stores, discounters and drugstores, the remaining $500.0 million (Hartman 2012).

[2] The online grocery market consists of "order and deliver," "order online and store pick-up," "specialty online," and "online order and shipment" sub-segments (Hartman 2012).

and consumers have less incentive to search among stores if they can more easily find a desirable product at a single retailer. However, if search costs are variable in the sense that considering each product involves a unique process of examination and consideration, then search costs rise in the number of products available.[3] In this study, our online data provide a relatively "clean" test of how search costs change with variety.

While the cost of searching among products may indeed be lower online than offline, online retailers are able to offer more variety, and more detailed information on each of the variants on offer. On its own, variety is well understood to be a source of market power (Boatwright and Nunes 2001; Borle, et al. 2004; Oppewal and Koelemeijer 2005; Richards and Hamilton 2006, 2014; Trindade 2010) due solely to consumers' preference for variety. Online retailers have other sources of pricing power, independent of variety. Degeratu, Wangaswamy, and Wu (2000) argue that online purchases are likely to be less elastic due to the importance of brand names, technology that facilitates habitual purchases, and, most importantly for our purposes, the volume of information available on non-price attributes. Although they do not test any of these explanations directly, they find the online price sensitivity of liquid detergent to be only 63% of its offline counterpart, and that non-sensory attributes such as fat content or perhaps sugar content, are important determinants of online sales. Similarly, Andrews and Currim (2004) and Chu, Chintagunta, and Cebollada (2008) find the demand elasticity for online grocery purchases to be far lower than the offline elasticity, and attribute the difference to higher incomes among online consumers, their demand for convenience,

[3] This debate is also related to the question of how choice behavior changes with variations in assortment. While McAlister and Pessemier (1982) argue that utility always rises in variety as each consumer has a higher probability of finding a product that matches his or her preferences, Diehl and Poyner (2010) find evidence to the contrary, namely that the probability of any choice at all falls as the number of choices rises. Some argue that such studies provide evidence of a "disutility of choice" while others maintain that the experiments used in these papers merely capture the effect of higher search and processing costs on the probability of choice.

buying larger baskets of items, greater brand loyalty, and having greater access to non-price attribute information. More non-price information allows online shoppers to more readily search for differentiated products. Variety, therefore, can have competing effects on retail prices as lower search costs may reduce retail prices, but may also facilitate higher retail margins. We attempt to answer this empirical question in our model of equilibrium price dispersion.

It is well understood that search costs are identified with price-only data. Hong and Shum (2006) estimate search costs using an empirical approach based on the mixed-strategy price equilibrium derived from Burdett and Judd (1983). By imposing the equilibrium conditions implied by both sequential and non-sequential search on the resulting price distribution, they are able to solve for the indifference points in the distribution of search costs.[4] Whereas Hong and Shum (2006) using an empirical likelihood approach, Moraga-Gonzalez and Wildenbeest (MGW, 2008) estimate the same type of model using more traditional maximum likelihood estimation. Wildenbeest (2011) extends the MGW approach to a differentiated-products environment in which firms compete directly in utility space (Armstrong and Vickers 2001), but focuses on single-product purchases (baskets of groceries). We use the maximum likelihood approach developed by Wildenbeest (2011), but base our model instead on the multiproduct mixed-strategy price equilibrium derived by McAfee (1995) that is relevant for firms that compete with multiple products. McAfee's model is more appropriate for the case where consumers do not necessarily buy multiple products from the same store, which would degenerate to a scaled version of the single-product problem, but one in which they costlessly receive information from multiple vendors of similar products. His model thus anticipates

[4] Although Hong and Shum (2006) estimate search costs for both sequential and non-sequential search, de Los Santos, Hortacsu and Wildenbeest (2012) reject the sequential search model, so we focus our attention on the non-sequential, or fixed sample size, alternative.

the era of internet shopping that we consider here.

The objective of this study is to determine the effect of product variety on search costs in a multiproduct, online, retail environment. We contribute to the literature on search and equilibrium price distributions in a number of ways. First, we estimate search costs in a multiproduct environment, not only multiple options within one category, but search within multiple categories on a single shopping trip. This environment is both realistic and likely revealing as it presents a significantly more complex decision space for individual consumers. Second, we extend the existing econometric methods of identifying price distributions for single products into a multiproduct framework. While Hong and Shum (2006), Moraga-Gonzalez and Wildenbeest (2008), and Wildenbeest (2011) develop maximum likelihood methods of estimating search costs for single products, the multiproduct problem is at once both more descriptive of actual shopping behavior, and inherently more complex. Third, we examine the effect of product differentiation on equilibrium price distributions, and search costs. Much of the previous work in this area uses data on homogeneous products in order to control for all sources of price variation other than variation in search intensity, whereas we extend the differentiated-products approach of Wildenbeest (2011) in accounting for product differentiation not at the level of the store, but for individual brands where choices are made. Adding product differentiation to the problem is necessary to form conclusions relevant to the retailing industry, and to examine the implications of search in a true differentiated, multiproduct environment.

In the next section, we provide some background on existent models of multiproduct search. In the third section, we describe our online grocery data, and highlight some observations that point to the salience of our research question, namely that online prices tend to be highly disperse and, evidently, less than perfectly competitive. In the fourth section, we develop an equilibrium model of price dispersion in which costly consumer search plays a prominent role, and emphasize the importance of multiproduct search in an online-retailing context. In this section, we also show how we incorporate the effect of variety on empirical search costs, both in single-product and multiproduct search environments. We present and interpret our estimation results in the fifth section, while the sixth concludes and offers more general implications for other forms of online retailing.

2 Background on Multiproduct Search

Models of online price dispersion typically consider search for a single product, books or electronics, for example. However, search costs in a multiproduct environment are likely to differ fundamentally from those incurred when only one product is searched, but the literature is unsettled on what this means for equilibrium retail prices. For example, correlation among prices is one of the key points of contention in the multiproduct search literature. Rhodes (2012) argues that non-advertised and advertised (shopped and non-shopped) prices will be positively correlated because retailers use advertised prices to attract customers to the store, and then must lower prices on the remaining goods because the arriving consumers are inherently price-sensitive. However, his prediction contrasts with Lal and Matutes (1994), McAfee (1995), and Hosken and Reiffen (2007) because none of these models assume that retailers reduce the price on other products once consumers are in the store.[5]

In fact, the opposite case seems equally plausible. Hess and Gerstner (1987) and Lal and Matutes (1994) argue that prices in a model of multproduct search should be negatively correlated, while McAfee (1995) and Hosken and Reiffen (2007) do not comment on the likely

[5] We consider only equilibrium pricing models, although others consider multiproduct search purely from a consumer perspective (Burdett and Malueg 1981; Carlson and McAfee 1984; Anglin and Baye 1987; Anglin 1990).

correlation of prices, and do rely on price-transparency at all. In these models, heterogeneity in valuation drives the mixed strategy equilibrium, not information on product prices. In equilibrium, firms manage the tension between charging higher prices to consumers who do not shop, and setting lower prices to attract those who do. Zhou (2012) considers a sequential model of search in which consumers incur a single search cost to learn the prices, and attributes, of several products at once. Search is sequential, products are horizontally differentiated by firm, each firm offers all products, and consumers search not only for price, but for attribute matches. The author reveals a different dynamic that separates single-product from multiproduct search, namely, because there are economies of scope in search, retailers can sell more items to a searching consumer by reducing the price of other products, in addition to the one that attracts the consumer. As in Rhodes (2012), therefore, products are priced like complements even though no prior assumptions are made regarding product substitutability, or complementarity. Of the counter-intuitive predictions, retail prices can actually fall with higher search costs in his model – higher search costs strengthen the economies-of-scope involved in searching, which strengthens the complementarity effect among goods that can be purchased together.

If the theoretical research was settled on how multiproduct and single-product search costs should differ, our empirical task would be simply to confirm or refute the received wisdom. However, based on the existing theoretical research, the relationship between search costs in a single-product and multiproduct world is an empirical question. In this paper, we seek to resolve that question by applying a simple, tractable model of search costs using price-only data. While Wildenbeest assumes consumers search for "shopping baskets" based on average prices, average valuations, and average utilities, retailing is inherently a multiproduct service (Rhodes 2012) so it seems natural to consider a multiproduct search process. McAfee

(1995) extends the mixed strategy equilibrium of Varian (1980) and Burdett and Judd (1993) in order to account for the fact that products sold at retail are often searched, and purchased, many at a time. In this case, it is not only the draw from one utility distribution that is important, but a single draw from a multivariate utility distribution. We show how the equilibrium utility distribution method of Wildenbeest (2011) can be extended to account for the effect of variety and search cost in a multiproduct retail environment.

3 Data Sources and Summary Observations

Our data are from Profitero.com. Profitero is a data-aggregation company that specializes in "scraping" retail pricing data from websites maintained by online retailers. Although Profitero gathers data from a wide variety of online retailers – retailers that sell goods from books to clothing and electronics – we focus on those that sell food online. Among global online food markets, the UK is arguably the most mature as all of the major retail players have a substantial online presence. Therefore, our data describes online grocery prices in the UK.[6] Tesco, Asda, Sainsbury's and Waitrose each offer online grocery purchasing and delivery service, while currently only Safeway offers on a nation-wide basis in the US. Fully 6% of grocery purchases are online in the UK, compared to less than 0.5% in the US in 2012 (Butler 2014). If price dispersion is indeed a competitive phenomenon, then the UK online food market represents an opportunity to test our hypotheses regarding the source and nature of retail price variation in online markets.

By comparing prices for the same product across stores in an online environment, we describe retail price distributions for food products, in the absence of traditional search

[6] Although our data analysis concentrates on the UK grocery market, current trends in online grocery sales in the UK suggest that insights gained in examining this more mature market may portend similar developments in the US.

frictions, in a non-parametric way. Estimating equilibrium price distributions using conventional scanner data for brick-and-mortar stores is problematic for a number of reasons. First, online shopping data are used frequently in the search literature because search costs estimated from brick-and-mortar shopping behavior include many elements that are not part of the "pure cognition cost" of finding, studying, considering and evaluating the prices and attributes for alternative products. Other than these pure costs, online shopping is nearly costless. Second, unlike conventional scanner data, the Profitero data does not suffer from aggregation bias. In most retail chains, local managers have substantial latitude to meet competitors' prices, which means that chain-wide price reports, even for one regional market, do not necessarily describe a single pricing strategy. Third, and perhaps most importantly, online food purchases represent a growing share of the grocery trade and look to become a significant part of the food distribution channel in the future. Moreover, developing an understanding of consumer search in online markets provides insight into brick-and-mortar retail behavior by revealing changes in the structure and performance of food markets following a decline in consumer search costs. Indeed, it is possible to price-check products from within a brick-and-mortar retail store using a cell phone.

Consistent with our objective to examine the cost of searching among differentiated products, we focus our attention on the top brands from a small set of focal categories. Specifically, our data consists of weekly price observations from 5 frequently-purchased categories: Breakfast cereal, yogurt, carbonated soft drinks, coffee, and jam. Each of these categorires are in the top 20 volume categories for all UK grocery sales, and are well-represented by brands that are likely to appear in all stores. In order to compare like products across stores, we choose 10 items from each category (6 in the jam category) that are recognized as the most popular brands by Marketing Magazine, and that are sold in each retailer for all 66 weeks (January 1, 2013 - March 31, 2014) of the data period. Table 1a shows the list of items in each category, which are matched across retailers by European Article Number (EAN) identifiers. In table 1a, the binding nature of the constraint that EANs match across retailers is notable. In fact, many of the most common brands described by Marketing Magazine are sold in different packages across retailers. The fact that these common brands are sold by some of the largest food manufactures in Europe points to both the capital required to produce retailer-specific packages, and perhaps also to the level of sophistication involved in recognizing the apparent benefits of obscuring search across retailers.

[table 1a in here]

In assembling this set of brands in each category, we do not intend to describe a typical UK shopping basket as in Wildenbeest (2011). Rather, identifying true search costs requires that we compare identical items across stores, and estimate the cost of searching among these brands. Of course, if we are to explain price dispersion among like brands as a function of search costs, it is first necessary to establish that price dispersion does, in fact, exist in our data. For this purpose, we calculate the mean and standard deviation of price and variety (measured as the number of stock-keeping-units (SKUs)), by category and retailer, and show the results in table 1b. This table reveals a number of observations that are important to our subsequent empirical analysis. First, there is substantial variation in both prices and number of items offered across retailers. In the empirical analysis below, we establish that the extent of price variation is indeed statistically significant at the brand level. Second, there does not appear to be an obvious relationship between variety and category-average price. The fact that no distinct pattern emerges is perhaps to be expected as we identify two possible effects, each counteracting the other, and flowing indirectly through search costs to the retail price. Although Asda offers the lowest category-average price in each case, there is no clear low-price option among the other three. Given that Asda also offers the least variety in two of the five categories, and nearly in a third, this is limited evidence of a price / variety trade-off that is apparent prior to estimation. Sainsbury's also appears to be a high-variety, high-price option, but this evidence does not rise to the level of a stylized fact as Waitrose often has higher prices. Third, because we include identical products in each category, the size of the price differences among online retailers suggests that there are clear incentives to search. For example, in the cereal category alone, the difference in price between the highest (Waitrose) and lowest (Asda) category-average price is fully 9%. If shoppers are inclined to search online for price alone, this price difference is likely enough to induce additional search. In summary, it appears as though our sample is sufficient to capture price dispersion in online multiproduct markets, and to estimate implied cost of searching among multiple differentiated products.

[table 1b in here]

4 Empirical Model of Search Costs with Price Data

In this section, we describe an empirical model that is used to estimate how search costs change in the depth of retail assortments, and how retailers endogenize assortments to influence consumers search behavior. Because we estimate this stage of the model using only price data, it is not possible to examine the effect of horizontal differentiation among products or retailers; however, we follow Wildenbeest (2011) in demonstrating how vertical differentiation can be incorporated by assuming retailers compete directly in utility and not in prices (Armstrong and Vickers 2001). Competition in utility is both valuable from an empirical perspective, and intuitive as retailers attract customers not only through low prices, but by offering superior service, keeping their stores clean, locating in convenient places, and stocking high-quality fresh foods. We then extend the Armstrong-Vickers-Wildenbeest approach to a multiproduct environment as in McAfee (1995) to demonstrate how search behavior in a more realistic retail environment differs, potentially, in a fundamental way.

We study price-only data for a number of reasons. First, price data are often the only data available for a particular cross section of products, retailers, or markets. In our case, there is no comprehensive data source for online price and volume data for food that covers the entire market. Profitero offers market-level data, but only prices. Second, there is a rich history of theoretical literature that examines price distributions as a means of explaining why the law of one price does not hold in reality. Stigler (1961), Varian (1980), Burdett and Judd (1983), Stahl (1989) and many others derive models of seller behavior that make strong predictions about how equilibrium price distributions behave, which demand empirical testing. Third, empirical researchers are able to extract remarkable insights from looking at price distributions only – insights sufficient to examine our core problem of how retail variety is related to search costs in online markets.

4.1 Search Costs in a Single-Product Environment

Using the price-only data provided by Profitero, we estimate a model of price dispersion among the four major online grocery vendors in the UK market. Because the items purchased are identical, we assume any price differences observed online reflect either vertical differentiation created by unobserved differences in customer service, delivery, payment or other aspect of perceived service quality, observed variation in retail variety, or costly consumer search. By controlling for vertical quality differences and assortment depth, we identify variation in search costs implied by the observed price differences, and are able to estimate how search costs vary with assortment depth. We extend our single-product analog to Wildenbeest (2011) to multiproduct price distributions in the next section.

Our estimation strategy extends Moraga-Gonzalez and Wildenbeest (2008) and Wildenbeest (2011) by incorporating variations in retail variety. Based on the literature cited above, search costs can either rise in variety (Diehl and Poynor 2010), or fall (McAlister and Pessemier 1982), so clearly the distribution of search costs must depend on variety in a fundamental way. We assume consumers use a non-sequential search strategy in which the number of products to search is determined by comparing the marginal cost of search with the marginal benefits of doing so, but where the cost is mediated by the number of products the consumer has to select his or her optimal consideration set from. In this sense, our empirical model is a test of the importance of confusion in brand selection – do retailers offer a wide range of products for the express purpose of making comparison more difficult? Such "strategic obfuscation" (Ellison and Ellison 2009) cannot be tested directly, but can form part of a reasoned explanation as to why retailers may find seemingly excessive variety to be, in fact, optimal.

Following Wildenbeest (2011), we define the utility available from purchasing from retailer $i = 1, 2, ..., N$ to be the difference between a retailer-specific valuation (w_i) and the price paid for an item sold by the retailer (we omit item indices for clarity): $u_i = w_i - p_i$. Provided consumers share a common valuation of quality, and firms produce quality using a constant-returns-to-scale technology, we write the valuation for a truly homogeneous product (h) as the difference between the retailer-specific valuation and the unit-cost of selling food (c_i): $w_i - c_i = h$. The retailer's margin is then the difference between the value of a particular product and total utility: $p_i - c_i = h - u_i$. Consumers search non-sequentially, or from a fixed sample of stores (De los Santos, Hortacsu and Wildenbeest 2012), and must pay a cost (s_i) to search for another product at another store. If the consumer searches j stores, therefore, the total cost of search is js_i. Search costs are random, drawn from a distribution

$F(s_i)$ with density $f(s_i)$.

We seek to determine how variety affects search costs, so we allow the unit search cost to depend on variety. There are (at least) two ways in which variety can affect search costs. First, with more products on the shelf, each product becomes more difficult to find as there are more alternatives to search through. Search costs may rise with variety if this effect dominates. Second, if the attribute space is more densely packed with products, the shopper will find product with his or her desired attribute profile more easily, that is, with less search. Search costs may fall as a result. Because we do not know *a priori* which effect dominates, the net effect is an empirical question. If the total number of products offered by retailer i is n_i, the distribution of search costs is given by $F(s_i(n_i))$, with density $f(s_i(n_i))$. Our objective is to estimate the parameters of this distribution.

Consumers optimize their search behavior in the sense that the marginal expected utility from searching j times should be at least equal to the marginal cost of doing so. Define the net marginal benefit of searching j stores as the difference in the expected maximum utility of searching j stores less the cost of search, or $nb_{i,j} = E[\max\{u_i\}] - js_i$ so in equilibrium $nb_{i,j} = nb_{i,j+1}$ and the cost of searching j stores is $s_{i,j} = E[\max\{u_{i+1}\}] - E[\max\{u_i\}]$. With this assumption, and the assumption that search costs are randomly distributed, the share of consumers searching j stores is given by:

$$\phi_j = \int_{s_{i,j}}^{s_{i,j-1}} f(s_i)ds_i = F(s_{i,j-1}) - F(s_{i,j}). \tag{1}$$

Following Armstrong and Vickers (2001) and Wildenbeest (2011), we assume firms compete in symmetric, mixed strategies in utility space. Let the utility distribution be denoted by $G(u_i)$ with density $g(u_i)$. A mixed strategy equilibrium implies that firms choose utility levels randomly anywhere in the support of the distribution of consumer heterogeneity, including one in which the utility is zero. Utility is zero for the set of consumers who search only once. If the utility for these consumers is zero, then the firms' margin is given by $h = p_i - c_i$, and the equilibrium profit is $\pi(0) = h\phi_1/N$, or the average valuation for a truly homogeneous product, multiplied by firm i's market share. More generally, the expected profit over all consumer segments, or consumers who search $i = 1, 2, ..., N$ times is given by:

$$\pi_i(u_i; G(u)) = (h - u_i) \sum_{j=1}^{N} \frac{j\phi_j}{N} G(u_i)^{j-1}, \tag{2}$$

or the retailer-specific margin multiplied by the share-weighted average demand for each store in the market. The mixed strategy equilibrium condition requires the expected profit from selling to consumers searching j times to equal the no-search profit, or:

$$h\phi_1/N = (h - u_i) \sum_{j=1}^{N} \frac{j\phi_j}{N} G(u_i)^{j-1}, \tag{3}$$

which is the condition we use to solve for the equilibrium distribution of utilities. According to Armstrong and Vickers (2001), firms compete in utility space, so maximizing expected profit with respect to u_i gives the first-order-condition:

$$\frac{\partial \pi_i}{\partial u_i} = \sum_{j=1}^{N} \frac{j\mu_j}{N} G(u)^{j-1} - (h - u) \sum_{j=1}^{N} \frac{j(j-1)\phi_j}{N} G(u)^{j-2} g(u) = 0, \tag{4}$$

where we have imposed symmetry in utilities subsequent to differentiation. Solving this expression for the density of utility, $g(u)$ yields

$$g(u) = \sum_{j=1}^{N} j\phi_j G(u)^{j-1} \left((h - u) \sum_{j=1}^{N} j(j-1)\phi_j G(u)^{j-2} \right)^{-1}, \tag{5}$$

which we then estimate using maximum likelihood over all observations for the parameters ϕ_j, $j = 1, 2, ..., N - 1$.

Estimating the model requires values for the utility implied by each store-choice. We recover the utilities in (5) by estimating hedonic pricing models with retailer-level fixed effects (δ_i): $p_{it} = \alpha + \delta_i + \varepsilon_{it}$ so that $w_i = \alpha + \delta_i$ and the utility from each retailer simply becomes the negative of the hedonic residuals (Wildenbeest 2011). We estimate the variety effect on search costs by allowing utility to depend on retailer-specific, and category-level variety as well. That is, we modify the hedonic pricing equation to include variety such that: $p_{it} = \alpha + \delta_i + V_i + \varepsilon_{it}$, where V_i is the variety, measured by the number of unique stock-keeping units (SKUs) in the category of interest for each retailer. Now, the valuation reflects consumers' preference for variety in each retailer, so the estimated utilities capture consumers' preference for variety. By comparing the estimated search costs with and without the variety effect, we are able to formally test the impact of variety on consumer search costs.

Once the distribution of utility is estimated, we then recover equilibrium search costs using the relationship captured in (1) subject to the equilibrium condition in (3). Specifically, the optimality of search costs from a consumer's perspective means that they can be written in terms of utility as:

$$s_{i,j} = \int_{u_n}^{u_m} (j+1)uG(u)^j g(u)du - \int_{u_n}^{u_m} juG(u)^{j-1} g(u)du, \tag{6}$$

where u_m is the maximum level of utility, and u_n is the minimum level. Defining $y = G(u)$ implies $dy = g(u)du$ so we eliminate the distribution of utility in (6) and write search costs solely in terms of y. Further, recognizing that the value of $G(u)$ at the maximum utility level is 1 and at the minimum utility is 0, we write (6) as:

$$s_{i,j} = \int_{0}^{1} u(y)([(j+1)y - j]y^{j-1}dy, \tag{7}$$

and the equilibrium level of utility is found by solving (3) for u as:

$$u = h \sum_{j=2}^{N} j\phi_j G(u)^{j-1} \left(\sum_{j=1}^{N} j\phi_j G(u)^{j-1} \right)^{-1}. \tag{8}$$

Numerically integrating (7) subject to the equilibrium condition in (8) provides values for the cost of searching across $j = 1, 2, ..., N$ stores.

With this model, we estimate the share of consumers searching either one, two, three, or all stores in the sample and recover estimates of actual search costs using the distribution function implied by (1) above. By estimating the implied search costs with and without the variety effect, we are able to recover the impact of variety on search costs. In this way, we test whether observed price dispersion among online retailers is due to true search costs, or is explained entirely by vertical differentiation among stores, and how the cost of searching among online retailers is affected by the variety offered in each. Indeed, the estimation routine developed by Moraga-Gonzalez and Wildenbeest (2008) and Wildenbeest (2011) summarized above provides a remarkably simple way to calculate the effect of variety on equilibrium search costs. Through calculating equilibrium search costs implied by utility distributions that do and do not include variety, and comparing the results, we infer the impact of variety on the cost of searching across multiple stores. For example, if the cost of searching a single store is higher for the model that does not include variety relative to the variety-inclusive model, we can conclude that variety reduces search costs, and reduces the incentive for consumers to consider searching across multiple stores. More consumers will search only a single store if they are better able to find products that meet their desired attribute set, and retailers are able to capture a greater share of the market. On the other hand, if variety serves as a means of strategic obfuscation (Ellison and Ellison 2009), greater variety will raise search costs. In equilibrium, consumers will search more stores as the perceived benefit of doing so is correspondingly higher, and each retailer will lose the most desired of shoppers: Those who do not consider more than one store.

4.2 Search Costs in Multiproduct Retailers

The model developed to this point is appropriate for estimating search costs when each of N retailers sells a single product. However, as McAfee (1995) explains, the search process in multiproduct retailing is fundamentally different.[7] If each product has a unique price distribution, then the mixed-strategy equilibrium involves a much more complicated, and often unsolvable, problem. In fact, he shows that there are many solutions to the equilibrium price distribution without additional structure. In this section, we describe a simple extension to the single-product approach outlined in the previous section to account for the fact that consumers evaluate many prices simultaneously in deciding the set of products to purchase.

While McAfee (1995) frames the multiproduct problem in terms of the marginal distributions of several products offered by the same retailer, so the formal development of the problem differs little from the single-product case, consumers are instead interested in the probability of drawing a set of prices together for products that they are likely to buy. They draw prices from a multivariate distribution from each retailer, and compare the costs of searching to the prospects of another multivariate draw from another retailer. Therefore, the extension to Wildenbeest (2011), extending the price-distribution logic to the distribution of utilities as we have done so far, is rather simple. Instead of facing the probability of a single utility from each retailer, the information that is most salient to shoppers is the probability of a particular set of utilities for the products they are interested in. We modify the approach above, therefore, by considering a multivariate distribution over utilities from

[7] Indeed, the scenario described by McAfee (1995, p. 84) in which "...consumers can costlessly visit all the stores..." is approximately the case for online grocery purchases so online shopping from multi-product retailers is "...qualitatively different than the single-good problem."

all five categories, rather than draws from univariate distributions for each item. The difference between our approach and Wildenbeest (2011) is subtle, but important. By considering a single basket of goods, Wildenbeest (2011) ignores the covariance of utilities among goods in the basket. Firms still compete by offering baskets of products, because this is the nature of multiproduct retailing, but when they choose a set of utilities with a particular probability, this probability takes into account the interaction of utilities among products. This extension is both conveniently simple, and powerful as it allows us to compare the implications of multiproduct search, and the impact of variety, directly with our findings from the single-product model.

More formally, the problem from the consumer's perspective now becomes one of comparing the marginal benefit of search to the marginal cost when the probability drawn represents a set of prices, which may or may not covary significantly with one another, to the cost of finding another set of prices. The share of consumers searching j stores now becomes:

$$\phi_j^C = \int\int \cdots \int m(\mathbf{s}_i) d\mathbf{s}_i = M(\mathbf{s}_{i,j-1}) - M(\mathbf{s}_{i,j}), \tag{9}$$

where $\mathbf{s}$ is now a vector of search costs associated with searching for items chosen from $c = 1, 2, 3, ...C$ categories within retailer i, ϕ_j^C is the share searching over the set of C products, and the multiple-integral is over the joint density of search costs $m(\mathbf{s}_i)$.

From the firm's perspective, they now offer a multivariate distribution of prices, and compete in mixed strategies, albeit in a more complicated way, just as they did in the previous problem. Firm's compete in utilities according to the Armstrong and Vickers (2001) assumption, so firms now choose the utility from each category, where utilities now represent draws from multivariate distributions $Q(\mathbf{u}_i)$:

$$\pi_i(u_i; Q(\mathbf{u})) = \sum_{c=1}^{C}(h_c - u_{ic})\sum_{j=1}^{N}\frac{j\phi_j}{N}Q(\mathbf{u}_i)^{j-1}, \tag{10}$$

which we solve in the same way as (4) above. Namely, the solution results in a system of equations analogous to (5) that we write as:

$$q_c(\mathbf{u}) = \sum_{j=1}^{N} j\phi_j Q(\mathbf{u})^{j-1}\left((h_c - u_c)\sum_{j=1}^{N} j(j-1)\phi_j Q(\mathbf{u})^{j-2}\right)^{-1}, \tag{11}$$

where $q_c(\mathbf{u})$ is the marginal density of utility from category c.

So the solution is formally very similar to the univariate case, but allowing for prices to covary with one another creates a solution that could be qualitatively different, and our solution is empirically tractable, unlike the theoretical models of McAfee (1995), Rhodes (2011), or Zhou (2012). That is, in the univariate case, the consumer forms expectations of the probability of a high or low utility from a particular retailer that derives from the attributes of, and price of, a single product. In the multivariate case, the consumer similarly forms expectations of the utility being offered, but that utility can be comprised of a highly desirable product offered at a low price, combined with a less desirable product, in another category, offered at a relatively high price, or any combination of price and attributes across multiple products. Any combination of positive or negative correlation among products – the key source of contention among recent theoretical studies on this issue (Rhodes 2011, Zhou 2012) – is possible, and can lead to dramatically different outcomes relative to the univariate case.

The key insight here is that we can summarize the outcome in our model with one utility draw, and an associated multivariate probability. Empirically, this problem is tractable, and we can compare the results directly to the single-product search case. As in the univariate solution above, we can also investigate the effect of changing variety simply by allowing utility for each product entering the multivariate problem to depend on the number of products offered by the retailer. If variety in one category at a specific retailer is positively correlated with variety and prices in another category in that retailer, then this will be reflected in the consumer's decision to search another store, or to remain and purchase from the store that changed variety. All we are interested in is the summary outcome, and the net effect on search costs and the probability of searching.

With the equilibrium vector of utilities in hand, we then calculate the share searching each number of stores (1, 2, 3, or 4 in our empirical application) and the implied costs of searching. Again, by calculating equilibrium utilities, and search costs, with and without variety effects, we are able to evaluate the effect of variety on search costs. In the next section, we describe the results associated with each hedonic estimation model, and the equilibrium search costs in the univariate, and multivariate search cases.

5 Results and Discussion

In this section, we first present the results obtained from estimating the utility associated with each brand in each category, with and without variety effects, and then the estimated search behavior in our univariate and multivariate search cases. We interpret our results in terms of both the significance of search in our focal product categories, and the revealed cost of searching within each. We conclude the section with a number of robustness tests that establish the validity of our empirical model.

In order to estimate search costs while controlling for the possibility that online retailers are vertically differentiated, we follow Wildenbeest (2011) and estimate hedonic pricing models for each category.[8] The results obtained by estimating the "without-variety" models

[8] Wildenbeest (2011) explains that interpreting the maximum observed price as the willingness to pay, and calculating each valuation as the difference between willingness to pay and the observed price, is another

appear in table 2. In this table, the item-specific terms represent the willingness to pay for each brand in the order listed in table 1a, holding the store constant, while the store effects show the average difference in willingness to pay relative to store 4 (Waitrose). In the cereal category, for example, each cereal can be expected to sell for an average of nearly 13 pence less in Tesco relative to Waitrose, almost 20 pence less in Asda, and over 5 pence less in Sainsbury's. Note that the store fixed-effect is not consistent in sign across categories. That is, soda sells for a premium relative to Waitrose in both Asda and Sainsbury's, while jam sells at a premium to Waitrose in Sainsbury's, but at a discount in Asda and Tesco. This is an important result as it shows that the perception of vertical differentiation cannot be assumed to be uniform across categories, so there is some feature of category-specific offerings that renders one store better than the others.

[table 2 in here]

One such factor may be the variety offered in each category. To investigate this possibility, and to model the variety-effect on utility, we estimate hedonic models inclusive of category-level variety, and report the results in table 3. In this table, the variety effects are predominantly, but not uniformly, positive. That is, the more variety in a category, the higher the willingness to pay for individual brands in each category and retailer. Although this finding is consistent with previous empirical results on variety effects on utility (Richards and Hamilton 2006), it is contrary to the evidence cited by Rhodes (2012) in motivating his theoretical model. Indeed, with respect to the cereal category, the estimates in table 3 reveal the extent of omitted variables bias likely present in table 2. Including category- and retailer-specific variety data causes the store 1 fixed effect to change sign. That is, although the estimates in table 2 imply that cereal sells at a discount to Waitrose in Tesco, once

alternative. We choose hedonic modeling as it allows us to easily estimate the effect of variety on product-specific valuations.

variety is included, it is apparent that cereal instead sells at a significant premium in Tesco. The fact that not all of the variety-effect estimates are positive points to the controversy referred to in the introduction. Namely, the choice overload hypothesis (Diehl and Poynor 2010) suggests that shoppers can become overwhelmed if the number of choices is too large, and fail to make a purchase at all. On the other hand, if greater variety implies that shoppers can find a brand that fits their ideal attribute combination more easily, the willingness to pay for individual brands in a store with more variety should be higher. In a reduced-form model such as this, we only see the net effect as either negative or positive. How these variety effects filter through to the cost of search, however, depends on the equilibrium impact on retailer pricing, and consumer search behavior.

[table 3 in here]

These estimates are shown in table 4 below. In the first 3 estimated columns of table 4, we show the share searching, the statistical significance of the share parameter, and the implied cost of search, and in the latter 3 columns, we show the same parameters with variety-effects included. Note in this table that we estimate only the first three share parameters as they are, logically, constrained to equal 1 in equilibrium. For the cereal category, we find that nearly 91% of shoppers consider only one store, while 7% shop two online stores, 2% shop three stores, and a negligible proportion shop all four stores. Meanwhile, the cost of search is substantial. Based on the average package prices reported in table 1b, the cost of searching one store is fully 2.56% of the shelf price. While this is proportionately lower than the search costs reported by Brynjolffson and Smith (2000) and Hong and Shum (2006) for non-grocery products, it is higher than the corresponding values reported by Wildenbeest (2011) in a similar context. Measured the same way, the other categories exhibit considerable heterogeneity in search costs, ranging from only 0.4% of the shelf price in the yogurt category to nearly 86% for coffee. While the estimate for coffee seems high, it is nonetheless only slightly higher than the relative price differentials or search costs reported in Brynjolffson and Smith (2000, 33%), Brynjolffson, Dick, and Smith (2010, 60%), and Hong and Shum (2006, 128% in one case). One pattern seems clear from these estimates – products that are more frequently purchased tend to have lower search costs, which is both intuitive and to be expected. These findings, however, tell us nothing about how variety affects search costs.

[table 4 in here]

For this purpose, we compare the estimates in the first three columns of table 4 to those in the final three columns. Beginning with the cereal category, we see that greater variety reduces the cost of search (0.0536 pence versus 0.0635 pence per package), and increases the proportion of shoppers who consider only one store – from 91% to almost 97%. This finding supports the argument that adding more products in a category increases the probability that a shopper is able to find a product that meets his or her desired attribute profile, reducing the cost of search, and the incentive to search another store for that ideal product. Although more variety means that the consumer must, potentially, search through more products to find one that is suitable, this effect is dominated by the ease with which the desired product is found. This finding is intuitive – if a consumer enters the cereal aisle looking for Special K Red Berries, and can only find Special K Original and Almond, then a decision must be made to purchase at the current store, or keep searching. This is a pure mental-processing cost that otherwise would not have had to happen. In the extreme case that cereal happens to be a must-have product, the consumer will leave the site and shop elsewhere. In the yogurt category, the cost of searching a single store falls only slightly, but the share searching only one store rises dramatically – from 85% to 97%. Examining individual elements of the cost-of-search function for the cause of this result, we find that the maximum utility value in the yogurt category rises substantially, increasing the value of search in the preferred store, without impacting search cost to as great an extent. Search costs in the soda category exhibit an even larger decline when variety effects are included – 22.4% lower than when variety is not included – and the proportion searching only one store rises in a manner very much like the cereal category. The similarity between the soda and cereal categories is not surprising because they resemble each other in structure, consisting of two major players that compete in variety and other non-price elements of the marketing mix.

Search costs and search activity follow a similar pattern in the remaining two categories (jam and coffee), namely that variety reduces search costs and causes the number of consumers searching only one store to rise. Variety thus appears to play a role that has not been considered elsewhere in the literature: By reducing search costs, retailers effectively reduce the incentives to search, and capture a greater number of consumers willing to pay local-monopoly prices in their store. Although the orthodox expectation is that online grocery shopping reduces search costs, and causes grocery markets to be more competitive, the outcome in the single-product search case is only partly consistent. Online retailers have the ability to offer far more variants of each brand than do bricks-and-mortar retailers, so search costs may fall online, but this does not necessarily mean that markets will be more competitive. These conclusions, however, are still based on a single-product analysis of search behavior. Whether multiproduct search differs is an empirical matter.

Recall that predictions of how multiproduct and single-product search differ in terms of their implications for retail prices is unsettled in the theoretical literature. While Zhou (2012) and Rhodes (2012) argue that consumers will search more intensively in a multiproduct environment, so retail prices will be positively correlated, moving both upward and downward together, the earlier literature (Lal and Matutes 1994; McAfee 1995; Hosken and Reiffen 2007) instead finds that retailers use low prices on some searched items to attract consumers to the store, raising prices on other items. In terms of the cost of search, we find a similar result to the single-product search case above (table 5). Namely, the cost of searching through products in a single store fall by an amount comparable to the single product-case (8.1%) when the variety effect is taken into account. However, search behavior differs dramatically, both with the multiproduct effect alone, and when the multiproduct search effect is combined with the variety effect. First, the results in table 5 show that consumers are more likely to search multiple stores when searching for products in many categories at the same time. This is perhaps to be expected as there is a chance the consumer will be unable to find his or her desired product in each category. In a single-product model, the probability of this one outcome alone drives the search to another store. In a multiproduct environment, however, if the consumer is unable to find the desired product in any of the five categories, she will likely try another store. As a result, fully 5.1% of consumers in the no-variety case and over 12.4% in the variety case are forced to search all four stores. More intensive search, in turn, suggests that retail prices across all categories are likely to be lower, supporting the "complementarity" argument of Rhodes (2012) and Zhou (2012).

[table 5 in here]

We next consider the variety effect in multiproduct search. The results in table 5 show that the complementarity outcome is sharpened when variation in variety is taken into account. Whereas the variety effect uniformly causes more consumers to search only one store in the single-product case, we see the opposite result when consumers search for products in all categories at once. Specifically, 78.4% of consumers search only one store without considering variety, while only 77.2% search one store when the variety effect is taken into account. More importantly, the fact that over twice the proportion of consumers search all stores in the with-variety case compared to the no-variety case suggests that variety does indeed increase search intensity, which potentially leads to lower retail prices, just as both basic theory and more complex models such as those developed by Rhodes (2012) and Zhou (2012) would have us believe.

Yet, variety still leads to lower search costs, so the precise mechanism in this case is not as clear. Recall the underlying structure of the search model: In equilibrium, consumers compare the marginal benefits of search to the marginal costs. In our empirical example, the cost of searching four stores relative to the cost of searching only one store averages 21.8% in the single-product / no-variety case, and 21.6% in the single-product / variety scenario. However, with search across multiple products, the ratio of four-store to one-store search costs changes from 22.1% without variety to 24.8% when variety is taken into account. Because the cost of search equals the benefits to search in equilibrium, there is clearly more benefit to searching all stores when variety is taken into consideration. This is, again, intuitive as consumers are likely to consider multiple stores primarily because of the different assortments offered in each. When searching for a specific flavor or package of jam in a basket that consists of the five categories considered here, a shopper is more likely to search across all four stores when the probability of finding a store that offers something specific and unique is higher. In the more general case, then, variety enhances the competitiveness of online markets rather than detracts from it as the single-category analysis implies.

There is a growing volume of research that predicts higher retail prices as online shopping becomes more prevalent. Bakos (1997), Anderson and Renault (1999), Degeratu, Rangaswamy, and Wu (2000), Cachon, Tierwiesch and Yu (2008), and Chu, Chintagunta, and Cebollada (2008) each argue for the potential rise in retail prices as consumers become less price sensitive – by whatever means – as they search for specific attribute mixes online. However, we find the opposite. If consumers search for multiple products on the same "shopping trip," or online session with the intention of buying groceries, we find instead that the ability of retailers to offer deeper assortments online can indeed make these markets more competitive than previously thought, or at least induce consumers to search more intensively. When searching for many products, across multiple categories, consumers are better able to find products with preferred attribute mixes, which lowers search costs. However, the marginal benefit of search also rises as the probability that another retailer offers the ideal product rises in the variety on offer. We find that the net effect results in more search across stores.

The primary implication of our findings is that online markets will be more competitive than the emerging literature would leave us to believe and, in fact, closer to what the common wisdom suggests. That said, we are unable to make a direct comparison between search online and off because our price-only data only reflects online transactions. Comparing our results for single- and multiproduct search suggests that the market power effects due to online search claimed in the literature may still hold true for retailers that sell items that are purchased infrequently and one at a time. Books, sporting equipment, electronic devices, or even shoes are prominent examples. Other retailers – supermarkets, drugstores, warehouse stores, for example – are designed to satisfy several needs at once. In these cases, the expansion of variety online is likely to represent a welfare gain to consumers, not just through the provision of more variety per se, but through the pro-competitive effects on online markets.

6 Conclusion

In this study, we investigate the empirical relationship between variety and search cost in an online environment in the context of an equilibrium, multiproduct model of search. With the rapid growth of online grocery retailing in a highly fragmented offline industry, how the structure of retailing will impact consumer prices is an issue of concern to not only public policy officials, but virtually all players in the supermarket industry. While orthodox economic thought suggests that lower costs of searching online will reduce prices, online retailing allows virtually unlimited variety, and variety is well understood to enhance retail market power, raising prices and retail margins. Which of these two effects dominates is the empirical question we seek to resolve here.

We estimate equilibrium price distributions with data from four major online supermarkets in the UK using maximum likelihood methods. We extend a well-understood empirical model of search costs to include variety-effects, and the effect of multiproduct search across five representative grocery categories. In a single-product search context, we find that greater variety reduces search costs, and increases the likelihood that consumers shop in only one store. Greater retail market power, and higher retail prices, ensue. However, when we consider an equilibrium model of multiproduct search, we still find that search costs fall in variety, but consumers shop across stores more intensively as the probability of finding one of their preferred products in another store rises. More intensive search has the potential to cause online grocery markets to be more competitive than markets for single-search goods such as books, electronics, or sporting goods. Although not surprising under conventional principles, this finding is contrary to much of the recent theoretical and empirical literature on the possibly disruptive effects of online retailing.

Our findings are subject to the caveat that we consider only prices from a set of online retailers. A more comprehensive analysis of this issue would compare search costs, variety, and multiproduct retailing between online and offline retailers. Further, we consider products from only five typical grocery categories. Because supermarkets, whether online or offline, commonly stock products from hundreds of categories, there is always the possibility that a different set of products would produce different results. We leave this question for future research, and validation of our results. Finally, our analysis is purely empirical. Multiproduct search models are notoriously complex, and often yield a multiplicity of equilibria (McAfee 1995). Future empirical research in this area would benefit from a more definitive theoretical background in order to test a more concrete set of hypotheses.

References

[1] Anderson, S. P. and R. Renault. 1999. “Pricing, Product Diversity and Search Costs: A Bertrand-Chamberlin-Diamond Model.” *RAND Journal of Economics* 30: 719-735.

[2] Andrews, R. L., and I. S. Currim. 2004. “Behavioural Differences Between Consumers Attracted to Shopping Online versus Traditional Supermarkets: Implications for Enterprise Design and Marketing Strategy.” *International Journal of Internet Marketing and Advertising* 1: 38-61.

[3] Bakos, Y. 1997. “Reducing Buyer Search Cost: Implications for Electronic Marketplaces.” *Management Science* 42: 1613–1630.

[4] Baye, M. R., J. Morgan, and P. Scholten. 2004. "Price Dispersion in the Small and the Large: Evidence From an Internet Price Comparison Site." *Journal of Industrial Economics* 52: 463-496.

[5] Boatwright, P., and J. C. Nunes. 2001. “Reducing Assortment: An Attribute-based Approach.” *Journal of Marketing* 65: 50–63.

[6] Borle, S., P. Boatwright, J. B. Kadane, J. C., Nunes, and G. Shmueli. 2005. “The Effect of Product Assortment Changes on Customer Retention.” *Marketing Science* 24: 616-622.

[7] Briesch, R. A., P. K. Chintagunta, and E. J. Fox. 2009. “How Does Assortment Affect Grocery Store Choice?” *Journal of Marketing Research* 46: 176-189.

[8] Bronnenberg, B. J., M. W. Kruger, and C. F. Mela. 2008. "Database Paper: The IRI Marketing Data Set." *Marketing Science* 27: 745-748.

[9] Brown, J. and A. Goolsbee. 2002. "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry." *Journal of Political Economy* 110: 481-507.

[10] Brynjolfsson, E. and M. Smith. 2000. "Frictionless Commerce? A Comparison of Internet and Conventional Retailers." *Management Science* 46: 563-585.

[11] Brynjolfsson, E., A. A. Dick and M. D. Smith. 2010. "A Nearly Perfect Market? Differentiation vs. Price in Consumer Choice." *Quantitative Marketing and Economics* 8: 1-33.

[12] Burdett, K. and Judd, K. 1983. "Equilibrium Price Dispersion." *Econometrica* 51: 955-969.

[13] Butler, S. 2014. "Grocers Rush to Open 'Dark Stores' as Online Food Shopping Expands." *The Guardian* (http://www.theguardian.com /business/2014/jan/06). Accessed on April 3, 2014.

[14] Cachon, G. P., C. Terwiesch and X. Yu. 2008. "On the Effects of Consumer Search and Firm Entry in a Multiproduct Competitive Model." *Marketing Science* 27: 461-473.

[15] Carlson, J. A. and R. P. McAfee. 1983. "Discrete Equilibrium Price Dispersion." *Journal of Political Economy* 91: 480-493.

[16] Carlson, J. A. and R. P. McAfee. 1984. "Joint Search for Several Goods." *Journal of Economic Theory* 32: 337-345.

[17] Chen, P., and L. Hitt. 2003. "Understanding Price Dispersion in Internet-Enabled Markets." Working paper, Carnegie Mellon University, Pittsburgh.

[18] Chevalier, J. and Goolsbee, A. 2003. "Measuring Prices and Price Competition Online: Amazon and Barnes and Noble." *Quantitative Marketing and Economics* 1: 203-222.

[19] Chu, J., P. Chintagunta, and J. Cebollada. 2008. "Research Note - A Comparison of Within-Household Price Sensitivity Across Online and Offline Channels." *Marketing Science* 27: 283-299.

[20] Clay, K., Krishnan, R., and Wolff, E. 2001. "Prices and Price Dispersion on the Web: Evidence from the Online Book Industry." *Journal of Industrial Economics* 49: 521-540.

[21] Degeratu, A. M., A. Rangaswamy, and J. Wu. 2000. "Consumer Choice Behavior in Online and Traditional Supermarkets: The Effects of Brand Name, Price, and Other Search Attributes." *International Journal of Research in Marketing* 17: 55–78.

[22] De Los Santos, B., A. Hortacsu and M. R. Wildenbeest. 2012. "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *American Economic Review* 102: 2955-2980.

[23] Dickson, P. and A. Sawyer. 1990. "The Price Knowledge and Search of Supermarket Shoppers." *Journal of Marketing* 54: 42-53.

[24] Goeree, M. S. 2008. "Limited Information and Advertising in the U.S. Personal Computer Industry." *Econometrica* 76: 1017-1074.

[25] Hess, J. D., and E. Gerstner. 1987. "Loss Leader Pricing and Rain Check Policy." *Marketing Science* 6: 358-374.

[26] Hong, H. and M. Shum. 2006. "Using Price Distributions to Estimate Search Costs." *RAND Journal of Economics* 37: 257-275.

[27] Hosken, D., and D. Reiffen. 2007. "Pricing Behavior of Multiproduct Retailers." *BE Journal of Theoretical Economics* 7: 1-43.

[28] Iyengar, S. S., and M. R. Lepper. 2000. "When Choice is Demotivating: Can One Desire too Much of a Good Thing?" *Journal of Personality and Social Psychology* 79: 995–1006.

[29] Jeuland, A. P. and C. Narasimhan. 1985. "Dealing - Temporary Price Cuts by Sellers as a Buyer Discrimination Mechanism." *Journal of Business* 58: 295-308.

[30] Kim, J. B., P. Albuquerque, and B. J. Bronnenberg. 2010. "Online Demand Under Limited Consumer Search." *Marketing Science* 29: 1001-1023.

[31] Koulayev, S. 2010. "Estimating Demand in Search Markets: The Case of Online Hotel Bookings (December 14, 2009). FRB of Boston Working Paper No. 09-16. (http://ssrn.com.ezproxy1.lib.asu.edu/abstract=1523926).

[32] Kuksov, D. 2004. "Buyer Search Costs and Endogenous Product Design." *Marketing Science* 23: 490–499.

[33] Kuksov, D. and J. M. Villas-Boas. 2010. "When More Alternatives Lead to Less Choice." *Marketing Science* 29: 507-524.

[34] Lal, R., and C. Matutes. 1994. "Retail Pricing and Advertising Strategies." *Journal of Business* 67: 345-370.

[35] McAfee, R. 1995. "Multiproduct Equilibrium Price Dispersion." *Journal of Economic Theory* 67: 83-105.

[36] McAlister, L., E. Pessemier. 1982. "Variety Seeking Behavior: An Interdisciplinary Review." *Journal of Consumer Research* 9: 311-322.

[37] Mehta, R., S. Rajiv and K. Srinivasan. 2003. "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation." *Marketing Science* 22: 58-84.

[38] Moraga-Gonzalez, J. L. and M. R. Wildenbeest. 2008. "Maximum Likelihood Estimation of Search Costs." *European Economic Review* 52: 820-848.

[39] Moraga-Gonzalez, J. L., Z. Sandor, and M. R. Wildenbeest. 2011. "Consumer Search and Prices in the Automobile Market." Working paper, Indiana University, Department of Economics. (http://www.kelley.iu.edu/mwildenb/searchautomobiles.pdf.)

[40] Oppewal, H. and K. Koelemeijer 2005. "More Choice is Better: Effects of Assortment Size and Composition on Assortment Evaluation." *International Journal of Research in Marketing* 22: 45-60.

[41] Pesendorfer, M. 2002. "Retail Sales: A Study of Pricing Behavior in Supermarkets." *Journal of Business* 75: 33-66.

[42] Rhodes, A. 2012. "Multiproduct Retailing." Working Paper, Department of Economics, Oxford University, Oxford, U.K.

[43] Richards, T. and S. Hamilton. 2006. "Rivalry in Price and Variety Among Supermarket Retailers." *American Journal of Agricultural Economics* 88: 710-726.

[44] Richards, T. and S. Hamilton. 2014. "Variety Pass-Through: An Examination of the Ready-to-Eat Cereal Market." forthcoming in the *Review of Economics and Statistics.*

[45] Stahl, D. 1989. "Oligopolistic Pricing with Sequential Consumer Search." *American Economic Review* 79: 700-712.

[46] Stahl, D. O. 1996. "Oligopolistic Pricing With Sequential Consumer Search and Heterogeneous Search Costs." *International Journal of Industrial Organization* 14: 243-268.

[47] Trindade, A. 2010. "Price and Variety in Supermarkets: Can Store Competition Hurt Consumers?" Working paper, Department of Economics, Northwestern University.

[48] Varian, H. 1980. "A Model of Sales." *American Economic Review* 70: 651-659.

[49] Villas-Boas, J. M. 1995. "Models of Competitive Price Promotions: Some Empirical Evidence From the Coffee and Saltine Crackers Markets." *Journal of Economics and Management Science* 4: 85-107.

[50] Wilde, L. L., and A. Schwartz. 1979. "Equilibrium Comparison Shopping." *Review of Economic Studies* 46: 543-553.

[51] Wildenbeest, M. R. 2011. "An Empirical Model of Search with Vertically Differentiated Products." *RAND Journal of Economics* 42: 729-757.

[52] Zhou, J. 2012. "Multiproduct Search." Working Paper, Stern School of Business, New York University, New York, NY.

Table 1a. Brands by Category

Cereal Brand	
Kelloggs Coco Pops 800G	Muller Light Gre
Kelloggs Cornflakes 750G	Activia Rhubarb
Kelloggs Crunchy Nut Bites 360G	Crunch Corner 5
Kelloggs Frosties 750G	Actimel Fat Free
Kelloggs Rice Krispies 700G	Yeo Valley Greek
Kelloggs Special K 550G	Onken Natural S
Nestle Cheerios 375G	Weight Watchers
Nestle Shreddies 750G	Shape 0% Fat St
Quaker Oat So Simple Org. Porridge 324G	Frobes Limited E
Weetabix 24Pk	Petit Filous Gree

Coffee Brand	
Carte Noire Ground Coffee 250G	Bonne Maman A
Kenco Pure Costa Rican 100G	Bonne Maman R
Taylors Rich Italian 227G	Rowse Clear Hon
Nescafe Original Coffee 200G	Rowse Squeezy C
Cafedirect Med. Roast Ground 227G	Rowse Squeezy L

Table 1b. Price and Variety Data by Item and Retailer

	Price		Variety	
	Mean	Std. Dev.	Mean	Std. Dev.
		Cereal		
Tesco	2.4863	0.4836	272.4960	42.7965
Asda	2.3538	0.3974	420.8991	53.5006
Sainsbury	2.5117	0.3583	1624.2160	232.1755
Waitrose	2.5671	0.4172	443.0076	76.9898
		Yogurt		
Tesco	3.7633	0.5811	152.6127	20.4722
Asda	3.6179	0.6113	157.6098	8.4738
Sainsbury	3.7592	0.7108	998.1110	112.8924
Waitrose	3.8281	0.5581	409.9381	6.0211
		Soda		
Tesco	2.2119	0.2747	105.3842	26.5682
Asda	2.1278	0.3749	95.9499	8.0242
Sainsbury	2.1721	0.4006	337.0887	54.0515
Waitrose	2.3118	0.3061	266.4560	17.0483
		Coffee		
Tesco	1.4843	0.3410	344.1006	74.0740
Asda	1.4590	0.4258	174.7514	11.0206
Sainsbury	1.5429	0.3743	2011.2330	103.9917
Waitrose	1.5031	0.3948	400.1559	44.1049
		Jam		
Tesco	2.2176	0.8430	159.0717	22.4901
Asda	1.8483	0.5736	362.4859	44.1759
Sainsbury	2.1641	0.8492	1475.7510	276.9420
Waitrose	2.1493	0.7647	424.9373	24.1799

Table 2. Hedonic Regressions by Item and Retailer

	Cereal		Yogurt		
	Estimate	t-ratio	Estimate	t-ratio	Estimat
Item 1	2.5185*	252.6118	2.1328*	252.4000	1.2277
Item 2	2.4772*	261.0306	1.9217*	226.8855	1.2111
Item 3	2.5797*	271.8335	3.3929*	399.1635	1.6990
Item 4	2.4615*	258.0178	2.6549*	314.1941	1.6957
Item 5	2.4006*	252.9621	1.7453*	206.5396	1.2254
Item 6	3.0120*	317.0516	1.2891*	152.3723	1.9522
Item 7	3.6919*	387.3987	1.5396*	182.2012	2.0017
Item 8	3.7995*	400.3635	1.4856*	175.8071	2.2198
Item 9	2.9940*	306.7582	1.9415*	229.7609	1.1192
Item 10	2.8270*	297.8936	1.5507*	183.2920	1.1762
Store 1	-0.1289*	-17.1611	-0.0131*	-1.9698	-0.0689
Store 2	-0.1967*	-26.1864	-0.3245*	-48.9457	0.0176
Store 3	-0.0553*	-7.3648	-0.0360*	-5.4359	0.0155
R^2	0.6574		0.7914		0.702

Table 3. Variety Effects by Item and Retailer

	Cereal		Yogurt		
	Estimate	t-ratio	Estimate	t-ratio	Estima
Item 1	2.5321*	78.6601	2.2381*	30.3055	1.4378
Item 2	2.4900*	78.0577	2.0270*	27.4439	1.4208
Item 3	2.5926*	81.2718	3.4982*	47.3492	1.9086
Item 4	2.4744*	77.4948	2.7601*	37.3699	1.9053
Item 5	2.4135*	75.6574	1.8505*	25.0578	1.4351
Item 6	3.0249*	94.7941	1.3943*	18.8781	2.1619
Item 7	3.7048*	116.0661	1.6449*	22.2739	2.2113
Item 8	3.8123*	119.5085	1.5916*	21.5319	2.4295
Item 9	3.0070*	93.7920	2.0468*	27.7158	1.3289
Item 10	2.8400*	89.0266	1.6559*	22.4198	1.3858
Store 1	0.1496*	3.2370	-0.2239*	-3.1145	-0.5514
Store 2	-0.2981*	-5.7340	-0.2272*	-2.6234	-0.6380
Store 3	-0.2409*	-5.0518	-0.1395	-1.7996	-0.1779
Variety 1	0.0011*	8.9167	0.0007*	3.9412	-0.0006
Variety 2	0.0002*	2.1300	0.0006*	5.6000	0.0016
Variety 3	0.0001	4.9661	0.0001*	2.9586	0.0002

Table 4. Cost of Search with and without Variety Effects

		Without Variety			With Variety		
Category	# Stores	Share Searching	t-ratio	Cost of Search	Share Searching	t-ratio	Cost of Search
Cereal	1	0.9096*	674.9215	0.0635	0.9697*	1,526.3625	0.0536
	2	0.0705*	14.4517	0.0324	0.0232*	12.9553	0.0281
	3	0.0199*	5.9105	0.0197	0.0071*	5.4856	0.0174
	4	0.0001		0.0132	0.0000		0.0119
LLF		-508.05			-8,080.57		
Yogurt	1	0.8488*	12,318.5714	0.0163	0.9708*	5.2547	0.0162
	2	0.0529*	255.0301	0.0094	0.0222*	3.1651	0.0085
	3	0.0148*	66.4736	0.0062	0.0070*	4.3549	0.0053
	4	0.0836		0.0043	0.0001		0.0036
LLF		-10,651.99			-23,389.86		
Soda	1	0.8934*	2.8892	0.4320	0.9698*	2.3022	0.3353
	2	0.0864*	2.4443	0.2175	0.0231*	2.1716	0.1761
	3	0.0203*	5.5816	0.1310	0.0071*	7.0138	0.1090
	4	0.0003		0.0875	0.0000		0.0742
LLF		-17,022.01			22,905.15		
Coffee	1	0.8761*	5.0265	1.2847	0.9501*	4.4192	1.2480
	2	0.1021*	4.9679	0.6395	0.0418*	4.5810	0.6390
	3	0.0218*	3.6692	0.3821	0.0081*	3.6675	0.3892
	4	0.0001		0.2539	0.0000		0.2622
LLF		-51,835.14			-41,069.79		
Jam	1	0.9036	1.2373	0.4567	0.9661	0.9902	0.4507
	2	0.0763	1.1344	0.2317	0.0266	1.0835	0.2353
	3	0.0201*	5.4331	0.1401	0.0073*	7.2928	0.1450
	4	0.0000		0.0939	0.0000		0.0985
LLF		-14,850.91			-14,817.88		

Note: A single asterisk indicates significance at a 5% level. LLF is the log-likelihood function value.

Table 5. Cost of Search with and without Variety in Multiproduct Retailing

	Without Variety			With Variety		
# Stores	Share Searching	t-ratio	Cost of Search	Share Searching	t-ratio	Cost of Search
1	0.7840*	674.9215	1.7853	0.7723*	1,526.3625	1.6410
2	0.1124*	14.4524	0.9423	0.0614*	12.9553	0.9243
3	0.0526*	5.9105	0.5819	0.0422*	5.4856	0.5893
4	0.0510		0.3945	0.1242		0.4063
LLF	-3.1552			-3.3750		

Note: A single asterisk indicates significance at a 5% level.