Shopping cost and brand exploration in online grocery *

Andrea Pozzi[†]

August 5, 2010

Abstract

This paper studies differences in consumers' brand exploration -the purchase of a brand not tried in the past- when shopping occurs online as opposed to in a brick-and-mortar store. Exploiting a new scanner dataset I can compare cross-channel behavior of the same households, shopping at the same chain, for identical items and identical prices. I document that brand exploration is systematically more prevalent in-store than online. My model quantifies the role of characteristic features of the e-commerce experience (existence of lists of favorites, difficult quality verification) in determining the result. Scarcer brand exploration online implies, in contrast to the conventional wisdom, that barriers for new entrants can be higher on the Internet channel. Counterfactual exercises suggest that new forms of online advertising could reverse this result and make the online channel more competitive.

*I am deeply indebted to Liran Einav, Tim Bresnahan, and Jakub Kastl for their advice and guidance. I also benefited from comments and suggestions by Paul David, Wesley Hartmann, Jon Levin, Sridhar Moorthy, Andres Musalem, Harikesh Nair, Sridhar Narayanan, Frank Wolak and participants in the IIOC 2009, the 7th ZEW Conference on the Economics of ICT, the EARIE '09 and in seminars at several universities. I thank the managers of the supermarket chain for their help in accessing the data and for sharing with me their knowledge of the industry and of their business practices. I gratefully acknowledge financial support from SIEPR, in the form of the Schultz Graduate Student Fellowship and of the B.F. Haley and E.S. Shaw dissertation fellowship, and from the NET Institute (http://www.netinst.org/). All errors are mine.

[†]Einaudi Institute for Economics and Finance, andrea.pozzi@eief.it

1 Introduction

Brand exploration -the purchase of a brand not tried in the past- is an important phenomenon in retail markets. Learning the extent to which it is possible to induce consumers to explore products they have not purchased in the past is key for manufacturers, as it informs their pricing (Klemperer, 1995) and marketing strategies (Bayus, 1992). It also carries policy implications given that markets where brand exploration is low have higher *de facto* barriers to entry (Schmalensee, 1974). In this paper, I study how the choice of the shopping channel (online or brick-and-mortar store) affects agents' propensity to perform brand exploration.

I show that customers tend to explore less when they shop online, and estimate a structural model that identifies and quantifies the importance of the drivers of such result. My findings imply that not only are firms able to exploit more their customer base on the online channel but they also face a considerably lower threat of successful entry from outsiders. This suggests that the role played by the online distribution channel in an industry should be taken into account when formulating both business strategies and policy analysis.

The traditional characterization of the Internet depicts it as a technology that reduces the cost of search. Tools such as shopbots and price search engines make it easy to gather information on alternative products and compare them (Bakos, 1997; Brynjolfsson and Smith, 2000; Clemons, Hann, and Hitt, 2002). More search means higher probability of trying a new item. Therefore, we should expect e-commerce to foster exploration and competition (Brown and Goolsbee, 2002).

On the other hand, it is increasingly recognized that differences between the Internet and brick-and-mortar setting extend past the level of search costs. For example, the online shopper does not have the chance to physically inspect the product which makes it harder to verify quality of online purchases. Furthermore, e-shopping websites present features meant to affect consumers' choices in ways that cannot be mimicked in a store. E-retailers provide recommendations based on customers profiling or allow to create personalized list of favorite goods. We expect these features to have an impact on product exploration and, as a consequence, affect the competitive environment. Although the outcome will depend on the website design of choice for each retailer, these practices are of general interests: they are becoming ubiquitous and firms are investing in perfecting them.[1] However, not much is known on the quantitative impact of these features on consumers' choice.

Identifying the effect on the online shopping setting on behavior is a challenging task because of the many potential confounding factors. First of all, the choice of becoming an online shopper is endogenous. To disentangle the effect of the channel from self-selection we would need to know how would an Internet customer behave if she were to shop in a brick-and-mortar store. Additional complications come from different pricing policies across the two channels (goods are typically cheaper online) and in the composition of the supply side (online there are more new entrants with none or little brand name).

My study explores a growing online shopping market, that of groceries (Goettler and Clay, 2006). I assemble new scanner data containing two years of grocery purchases by a panel of some 11,000 households who shop *both* online and in-store at the same supermarket chain. In-store purchases are tracked through usage of a loyalty card which is also required when the customer registers for the online service on the grocer's website. This allows me to link in-store and online trips for the same household.

A number of features of the setting at hand make it particularly attractive to isolate the effect of shopping online and allow for an improvement over similar attempts in the literature. First, I observe the same household switching back and forth between the two channels; whereas most previous studies compared a sample of online shoppers with a different sample of traditional shoppers.[2] This considerably reduce concerns related to sample selection. Moreover, both in-store and online purchases occur at the same supermarket chain. This ensures that the set of brands carried, prices and promotions, and loyalty to the retail chain are all the same in both channels. If I was simply trying to compare trips at a traditional grocery store with purchases made at a different online retailer, each of these could have

[1] In 2006 Netflix, an online video rental company, offered a \$1M prize to those who could achieve a substantial improvement on their algorithm for rental recommendations. Contestant in the challenge were even granted access to proprietary data on previous customer rentals.

[2] Exceptions are Chu, Chintagunta, and Cebollada (2008) and Brynjolfsson, Hu, and Simester (2006) who also use a panel of cross channel purchases, but with different focus.

confounded the analysis.

I focus my analysis on the breakfast cereals product category. This category is a classic example of niche filling strategy (Scherer, 1979; Schmalensee, 1978) and we observe an extremely large number of differentiated brands marketed by each manufacturer. As a consequence, there is large scope for brand exploration which is the economic phenomenon I study in this paper. I start by documenting that households in my sample are almost 8 percentage points more likely to try a cereal brand they did not purchase before when they are shopping in a brick and mortar store than when they are purchasing groceries online. This stylized fact is robust to different specifications and is consistent with previous findings.[3] My aim in the rest of the paper is then to understand the causal mechanisms behind this descriptive result. I point to three different potential drivers of the pattern seen in the data.

First, the possibility that households sort their transactions has to be taken into account. Even though selection is less of a concern in my study since everybody shops online and offline, the decision of when to shop online or in a store is likely endogenous. Purchasing grocery online is a time efficient way of shopping. This suggests that a customer might be more likely to choose the online channel when she is short of time, a situation in which she may also be less likely to explore new brands.

Next, I consider features of the Internet environment that both reduce the cost of shopping and make the online shopping experience less flexible. Examples of such features are lists of favorite or recommended items. In my application, the grocer's website offers the option to save on the time spent browsing, by choosing to purchase from a "one click" list of items already selected by customers in their past shopping trips. This cost reduction comes at the expenses of their freedom to browse the virtual aisles and -potentially- spot new brands. By construction, this lowers the amount of online exploration. The effect of features similar to this *past shopping history* list has been overlooked by the literature on e-commerce, despite evidence that the cost of browsing a webpage or typing can be substantial (Hann

[3]Degeratu, Rangaswamy, and Wu (2000) use data from Peapod and observe that brand switching is lower in Peapod orders than in brick and mortar stores. Danaher, Wilson, and Davis (2003), with data from a grocery chain that offers also online delivery, find that brand loyalty is higher when customers choose to purchase on the Internet.

and Terwiesch, 2003). For more traditional online goods (e.g. airplane tickets or electronics) which are seldom purchased and expensive, this is unlikely to be a relevant issue. However, in the case of a repeated activity involving relatively cheap items, like grocery goods, the appeal of a reduction in the shopping cost can be much greater.

Finally, quality verification online is harder than it is in-store (Bhatnagar, Misra, and Rao, 2000; Jin and Kato, 2007). The information content of online search is poorer than in the in-store setting. Exploration of new brands is the search for items with more desirable attributes than those currently in the consumer basket. If assessing the attributes of potential trials is more difficult online, the probability of exploration being successful is lower for online purchases. As a result, taste for new brands will be lower for the same consumer when she shops on the Internet.

The importance of these effects is quantified by estimating a structural model of consumer behavior. In the model, consumers first select the channel where they want to shop and, conditional on that choice, pick a cereal brand. Results indicate that the sorting element plays a role. Residuals from the channel selection equation and unobserved shocks to the taste for unknown brands are negatively correlated. This means that specific circumstances that make it more likely for an agent to go online (e.g. lack of time) make her less likely to try new brands as well. Search for shopping cost reduction through the *past shopping history* list turns out to be a very important factor in explaining why consumers avoid exploring on the Internet. The estimates imply that the benefit of not having to browse the website -by resorting to the *past shopping history* list- is worth about $4 to the consumer.

The counterfactual exercises provide additional insights on the economic significance of my results. Altering the design of the website to remove the *past shopping history list* boosts brand exploration by 23% on the Internet channel. I furthermore show that a new entrant's penetration is slower online. While the brand achieves 60% of its potential in-store market share within 18 months, it only reaches 40% of its target online market share. This finding gives a measure of the additional entry barriers that the online technology creates in the environment I study. At the same time, when I simulate the introduction of features that could push exploration of new brands online (like context ads or recommendations), I find that the entrant penetrates the market much more quickly on the Internet.

My results provide insights on the impact of the online shopping technology on a dimension of consumer demand behavior. This research ties in to a growing body of work taking into consideration the impact of popular features of e-commerce website. Prominent examples include Ellison and Ellison (2009) study of obfuscation techniques and Ackerberg, Hirano, and Shahriar (2009) analysis of the Buy-it-now option in e-Bay auctions. By showing that the Internet can make exploration of new brands relatively more costly online, generating an advantage for brands already popular and increasing entry barrier for outsiders; this paper also relates to work testing the "frictionless commerce" paradigm (Brynjolfsson and Smith, 2000; Chen and Hitt, 2002; Brynjolfsson, Dick, and Smith, 2010).

The rest of the paper is organized as follows. Section 2 presents the data and provides institutional details on online grocery business. In Section 3, I discuss descriptive results on the cross channel rate of trial of new brands. Section 4 develops the structural model of channel choice and demand for brands. The estimation strategy is explained in Section 5. Section 6 reports the results, and Section 7 presents the counterfactuals. Section 8 concludes.

2 Data and institutional background

Data for the analysis come from a large, national supermarket chain, which operates more than 1,500 store across the US and also offers the option of shopping for groceries online, through the company website, and receive them delivered at home. I observe scanner level data for each purchase made by the 11,640 households who shopped at least once in a supermarket store and at least once through the online service between June 2004 and June 2006. For each one of those 1,829,254 shopping trips I observe the date, the list of all the items purchased (as identified by their bar code or UPC), price, and quantity purchased. Usage of coupons and price promotions are also recorded in the data allowing to identify that price actually paid by the customer. Finally, I have information on whether the purchase took place in a brick-and-mortar store or if it was an online order. Customers are identified through their loyalty card number;[4] the chain is able to match cards belonging to different members of the same household under the same identifier.

The online service was first offered in 1999 but it was substantially re-organized in 2002. The option of buying online is available only in a limited number of metropolitan areas. Nevertheless, in areas reached by the service, the online distribution channel plays a non negligible role. In my sample, Internet trips account for 9% of the total number of trips and generate about 25% of the revenues. The online service has no separate warehouses; orders are fulfilled using stocks available in stores for each area covered by the service. For this reason, I do not expect the stockout process to have any systematic impact on differences between the two channels.[5] The service runs seven days a week, customers pay a delivery fee and can pick the delivery time, conditional on availability.

To access the online service, customers have to register by providing their address and phone number. Registration also requires to enter the loyalty card number; this allows me to link online transactions of a customer with her in-store ones in the data. The registration process is quick and easy and provides customers with a username and a password. Online, the customer can choose how to perform her shopping (Figure 1). As a first option, she can browse the online store, where aisles are represented as a series of nested links (Figure 2), and items can be displayed in alphabetical order or ordered by price. The alternative is to shop from the list of items bought during the last visit or in all her shopping history. Description of the items listed include characteristics, price and presence of discounts, a picture of the item, and information on nutrition factors. The retailer is committed to offer

[4] Purchases made by households not owning a loyalty card are not part of my data. This is not a big concern since the chain pushes usage of the loyalty card (that entitles to discounts and special offers on many items) and estimates that more than 85% of the customers hold one.
As for usage of the loyalty card, figures are also very high. Einav, Leibtag, and Nevo (2010) analyzes shopping trips reported by Homescan households for a particular retailer that tracks customers through loyalty cards. Less than 20% of the shopping trips recorded by a household in Homescan do not find a match in the retailer's data. This is an upper bound since the difference can be explained by trips where the card was not used but also by mistakes made by the household while recording the trip. For example, she can erroneously report the identifier of the visited store, or the date of the trip.

[5] Online customers are offered the possibility to specify instructions to be followed in case one item in their shopping list was not available. The three options offered are: "no substitution", "same size, different brand", and "same brand, different size".

goods at the same price online and in-store. Therefore, customers face identical prices in the two environments.[6]

Table 1 compares trip characteristics across channels. Online transactions are remarkably larger in total expenditure, size of the basket, and number of unique items purchased. This is driven by two factors. The first is the minimum $50 order rule to qualify for home delivery, that forces online orders to be worth at least that much. Moreover, customers tend to exploit the home delivery service by placing larger orders online and ordering systematically more heavy and bulky items (Chintagunta, Chu, and Cebollada, 2009; Pozzi, 2009). To make online and in-store orders more comparable, I condition on "large" trips in the right panel of the table. Large online trips are still bigger than large in store trips (and significantly so) but the magnitude of the difference is not as economically significant. Online and in-store trips appear to be comparable in terms of the most popular product categories purchased in each channel. Table 2 shows that 9 out of the top 10 categories purchased overlap in the two channels.

For the purpose of the present study, I focus on purchases of breakfast cereals. Breakfast cereals represent an excellent product category to analyze product choice and brand switching decisions by the households due to the large number of existing brands (more than 100 different brands with positive sales). Furthermore, it is a popular and frequently purchased category, featuring prominently both in online and in in-store sales (more than 140,000 trips). Finally, this choice eases the bias due to exploiting data from a single retailer as the grocery chain under analysis is a major player in this product category[7].

Whereas my economic object of interest is the brand, an observation in the data is a UPC code; Figure 3 explains the difference between the two. Each of the cereal manufacturers in my data distributes a number of different cereal brands. For example, Kellogg's produces Rice Krispies as well as Special K. Most cereal brands come in different varieties; for example

[6]Note that this does not imply that prices are the same in all the stores. The retailer's price strategy is based on price areas, and the online customer is offered prices matching those of the price area of her IP address.

[7]Based on a sample of Homescan data for 2004, the chain is among the top three retailers both for market share and number of trips in the breakfast cereal category in the US. If we restrict the analysis to markets where the chain is actually operating, it becomes by far the largest retailer in the product category.

it is possible to choose between Frosted Rice Krispies and Berries Rice Krispies. Finally, varieties are available in different box sizes. When I refer to a brand, I bundle together different box sizes and varieties: the small box of Frosted Rice Krispies, and the large box of Berries Rice Krispies to the purpose of this study are the same brand.

Table 3 reports some relevant descriptive statistics for trips to the supermarket involving a cereal purchase by distribution channels. The average size of the basket and the net expenditure are higher online, both reflecting the $50 minimum order constraint to access the online service and the incentive to stock-up in online orders, once the cost of the home delivery fee is sunk. However, once we focus on cereal purchasing behavior, the differences are less marked. The number of cereal boxes purchased per trip is just larger when the customer orders online; the gap between the number of different cereal brands purchased in an online vs. offline trip is even smaller.

For a random subsample of households in my data, the grocer provided information on the address, edited to prevent identification of the household. This information is available for 6,155 of the households who purchase breakfast cereals at least once. I match those households with demographic data from the Census 2000 at the block group level[8]. The demographic data contain information on the share of black and hispanic people in the block, the share of families, fraction of population with college degree, fraction of people employed, age of the head of the household, and income per capita. Table 4 provides an overview of the demographics for the households in my sample, as given by the blocks they live in.

[8]Matching at the block group level, rather than at the usual 5-digit zipcode level, has two main advantages. Block groups are smaller: their boundaries never cross county or state limits (as opposed to census tract boundaries), and are designed to include relatively homogeneous population. Hence, not only am I averaging demographic characteristics over a smaller set of people, but also over a set of people that is more likely to be similar.

3 Descriptive results

I start by documenting the relationship between the choice of the shopping channel for the grocery purchase and brand exploration. I define brand exploration as the purchase of a cereal brand that the household has not bought in her previous shopping history. Note that, under this definition, a new brand is intended as new to the buyer rather than only to the market. A cereal brand can have been around for a long time and still be a new brand to a household who never tried it. In order to assess whether a certain brand purchase represents a new trial, I need to construct the set of brands known to each household at t_0. I use the first three months of data (June 2004 to September 2004) to recover the initial shopping history of each consumer. This implicitly assumes that three months is a period of time long enough to observe the customer purchase all the brands she is already familiar with[9].

I observe 142,025 supermarket trips involving purchase of breakfast cereals, performed by 9,175 different households. In 52,461 trips the customer buys a cereal brand she never tried in her previous observed shopping history. In 42,957 transactions, more than a single brand of cereals is bought, allowing for multiple brand explorations in the same transaction. Considering multiple cereal purchases in the same shopping trip separately, I obtain 61,216 trials of new brands in 184,982 purchases.[10]

A look at the raw data (Table 5 top panel) suggests that the average amount of brand exploration in store is significantly higher than the same figure for online trips. The difference persists even when I consider only "large" trips, making online and offline trips more comparable. To assess whether the result is robust to the inclusion of controls, I estimate

[9]See the Appendix A1 for a validation of this assumption and for sensitivity analysis of the results to different initial conditions

[10]This implies a 33% probability of exploration for each trip. While the figure may seem high, it is in the range of other such estimates derived by scan data. Dube, Hitsch, and Rossi (2009) use 6 products (distinguishing between different size within a brand) of frozen orange juice and 4 brands of margarine and observe that probability of repeated purchase for a given brand oscillate between 77% and 90%. Shin, Misra, and Horsky (2007) restrict their analysis to 7 toothpaste brands (accounting for 86% of the market) whose probability of repeated purchase ranges between 46% and 57%. Shum (2004) adopts an approach very similar to the one of this study both in defining the brand as level of observation, and in considering a large number of them in the analysis (the top 50, with a cumulated market share of 75%). Moreover, his measure of brand loyalty is such that a brand switching event for a household in his sample is perfectly comparable to an exploration trip by one of my households. He finds that the probability of switching is around 50%.

the following probit model of trials:

$$Trial_{it} = \alpha + \beta Online_{it} + X_i\gamma + \varepsilon_{it} \tag{1}$$

where $Trial_{it}$ is a dummy variable that equals one if consumer i performs brand exploration in shopping trip t. $Online_{it}$ is also a dummy variable that denotes that the trip occurred online, X is a vector of consumer characteristics.

Results are reported in Table 6. Unconditionally, consumers are almost 13 percentage points less likely to try a new brand of cereals when they are shopping online. The result is robust to introducing demographic controls and day-of-the-week fixed effects. In the preferred specification I control for time-invariant unobserved heterogeneity with household fixed effects; the gap reduces to 7 percentage points but is still significant both statistically and economically.

3.1 Potential explanations for lower brand exploration online

The special features of my setting rule out sample selection, differences in price or quality, and reputation of the retailer as causes of the wedge between online and in-store behavior. Below I present other potential explanations.

Sorting of trips. The choice of ordering grocery on the Internet is endogenous and can be determined by factors that are also correlated with the probability of exploring a new brand. For example, if the customer views online shopping as a time saving technology, she will be more likely to order online when she feels under time pressure. Enlightening in this sense is the quote from a user of the online service collected by the grocer in a customer survey:

> *"I can't do without this service. It is a necessity for busy families."*

Being short of time is also a condition that does not favor experimentation. This would generate a spurious negative correlation between brand exploration and online shopping.

Figures 4 and 5 provide evidence consistent with this story. The first panel shows the fraction of in-store (online orders are excluded to avoid composition effects) trips in which the customer chose to perform exploration, relative to the total number of cereal purchases, for each day of the week. The second plot displays the share of online orders relative to the total number of shopping events, for each day of the week. The two series are negatively correlated. The amount of new brand trial spikes during weekends, when people are also likely to have more time to do their grocery shopping. At the same time, weekends feature the lowest share of online orders. This is consistent with the existence of an unobserved shock to the utility of shopping online correlated with an unobserved shock to the utility derived from switching to a new brand (taste for exploration).

Uncertainty over quality. While the consumer knows the utility she derives from consumption of a brand already tried, she can only have expectations about the quality of new brands (Erdem and Keane, 1996; Crawford and Shum, 2005). Uncertainty can lead the consumer to discount the utility she would derive from a switch. The reason for this cost lies in the nature of experimentation and the cost is not specific to the online environment. However, there is a gap in the possibility to inspect the product between the two channels. The store experience is richer, whereas online the customer can only read the nutritional information and look at a picture of the box. As a consequence, the consumer may be more worried about her judgement on the quality of new brands. This results in the cost of experimentation being higher online and could therefore explain the negative correlation between online shopping and brand exploration.[11]

Design of the website and shopping cost. Characteristics of the website design are known to affect consumers' behavior (Burke, Harlam, Kahn, and Lodish, 1992; Ellison and Ellison, 2009). Cutting edge e-commerce websites offer the possibility of creating lists of favorite or frequently purchased goods, provide recommendations and offer advertising through banners or pop-ups. All of these can impact on the propensity to try new brands. For instance, the option to shop from the list of items (UPCs) already bought in the past spares the customer the need to browse the website in search of the items she needs but prevents

[11] However, Mazar, Herrmann, and Johnson (2007) provide experimental evidence that i) customers evaluate fruit cereals based on non sensory attributes ii) they seem to be more accurate in matching their declared preferences when they shop for cereals online.

exploration of new items. Recommendations or pop-ups can generate the opposite effect, stimulating brand exploration. In my application, the grocer's website featured the option to shop from a "shopping history list" but did not offer context ads nor recommendations providing an opportunity to identify the effect of the former.

In my setting, both uncertainty over quality and website design tend bias customers towards purchasing known brands online. The two effect can, however, be separately identified exploiting the fact that the "shopping history list" includes UPC's and not just brands. Therefore, it should lead the customer not only to stick to previously purchased brand , but also to the specific variety or size of the box, within a brand. The latter would not occurr if the behavior were driven by uncertainty over brand quality. In Table 5 (bottom panel), I condition on instances in which the household decided to purchase a brand already bought in the past and look at whether persistence in purchase of the same UPC is different across channels. Indeed, the share of those who chose to purchase the same brand with exactly the same features (same box size, same variety) is much higher for online purchases.

Columns (5) and (6) in Table 6 reinforce this point by showing that online shoppers are around 10 percentage points less likely to try new varieties or box sizes of a known brand. I interpret this result as evidence that households seek to reduce their cost of shopping (by exploiting the *past shopping history* list) and this has implication for brand exploration. This explanation is consistent with the literature that measure the monetary cost of online search (Brynjolfsson, Dick, and Smith, 2010; Hong and Shum, 2006), reporting large numbers. It also squares with evidence that consumers with longer experience in Internet shopping tend to develop routines and cut on the length of their transactions (Johnson, Bellman, and Lohse, 2003). The logic of these arguments is reinforced if we put it in the context of a frequent activity, as is grocery shopping.

3.2 Additional explanations and robustness

Other dimensions of the difference between online and in-store shopping can contribute to explain the main result in Table 6. To begin with, there is a striking difference between the size of online and in-store trips, with the former being on average much larger. Column (2) in Table 7 suggests that this is not a factor in explaining difference in brand trial. The gap in exploration between the two channels remains once I control for the size of the trip. This is consistent with observations made on Table 5.

Online and traditional shopping channel also radically differ on characteristics and effectiveness of promotion strategies. This could drive my results since promotional activities are a very effective way to induce trials of new brands. A first issue concerns coupons whose usage is massive in the cereal product category (Nevo and Wolfram, 2002). The fact that paper coupons can only be redeemed in-store could be behind the higher propensity to new trials on the traditional channel. However, coupons do not play a major role in the supermarket chain under analysis: less than 2% of the transactions involve usage of paper coupons. The chain tries to foster usage of the loyalty card linking discounts to the membership card.[12] Moreover, usage of paper coupons is stronger among the elderly (Aguiar and Hurst, 2007) which represent only a small fraction of my sample (as we would expect, given that it only includes people who have tried the online shopping channel).

The specification in column (3) includes a dummy for the purchased item being on discount. Price promotions are highlighted both online and in-store by a red price tag and the icon of the loyalty card. The effectiveness and salience could differ across channels. Brands on promotion are more likely to be tried, as expected. The interaction between the promotion and the channel dummy is positive and significant: if anything, promotions are more effective in drawing exploration on the Internet. This suggests that a bias from not accounting for price promotions would go against my findings.

The two channels differ even more when we consider possibilities for promotional activity other than price discounts. The store setting offers the possibility of end-caps display or to offer free trials; whereas this is not possible online. To control for in-store promotion

[12] Prior to the beginning of my sample, issuing of coupon books had even been discontinued by the grocer and all the promotions were linked to usage of the loyalty card. Coupons books were later reintroduced but do not play a major role. To get an idea, I collected weekly booklets for a single price area between April and October 2008. Those booklets are mailed as general advertising and list price promotion for the week. They also contain few paper coupons that can be cut and used by the customer. Out of the 27 weekly booklets I collected, only three offered a coupon for a breakfast cereal brand.

of specific brands I include brands fixed effects (column (4)) and brand-week fixed effects (column (5)) since promotion generally take place on a weekly basis. In both cases the gap in exploration between the two shopping channel does not fade away.

Another alternative explanation revolves around the role played by kids. Kids are potential triggers of brand exploration in cereals, as they follow fashion and peers' example. If we assume that children accompany parents and influence them in the store but do not assist to Internet orders, this could explain the difference in the rate of brand exploration across the two channels. However, brand exploration is not limited to kids' cereal. The mean of the "new trial" dummy is .36 for kids cereal, only slightly above the overall figure of .34. Moreover, the gap between online and in-store rates of exploration is not dramatically more pronounced for kids' cereals than it is for other types of cereals. Online exploration in kids cereals is .28, while in-store is .39, an 11 percentage points gap. The gap for the sample without kids' cereals is 9 percentage points (.25 online vs. .34 in-store). I perform two robustness checks of the main descriptive exercise. In column 6 I exclude from the sample all the trips where kids' cereals were purchased to find that the qualitative results do not change. Finally, in column (7) I interact the online dummy with the census variable "family". This should give a measure of how much of the effect is truly driven by families showing up at different rates online and in-store. The interaction dummy is not significantly different from zero.

4 Model

In order to disentangle the impact of trip-sorting behavior from the causal impact of the online channel, I develop a model of consumer behavior. I assume that households face two sequential decisions. First, they have to select the channel where they want to shop. Each trip can take either of the two forms: a visit to a brick-and-mortar supermarket or an online order through the chain's website. Conditional on the choice of the channel, consumers select a brand of cereal.

I index each of the N consumers with i, each of the J UPCs with j, and each of the T_i trips made by customer i with t. With the notation Ω_{it} I refer to the set of cereal brands purchased in the past by consumer i, as of trip t. Finally, h_{it} indicates the set of UPCs purchased by consumer i prior to trip t.

4.1 Choice of the Shopping Channel

The utility from making trip t online for consumer i living in price area p is

$$z_{it}^{*} = \gamma_0 + \gamma_1 * Distance_i + \gamma_2 * Fee_{pt} + X_i \gamma_3 + \gamma_4 * Weekend_t + \mu_i + \theta_{it} \quad (2)$$

The regressors include variables relevant for the decision between shopping online and in-store. *Distance* measures the distance in miles between the house of the customer and the closest grocery store of the chain. *Fee* is the amount in dollars paid for home delivery in online orders: it ranges from 0 (free delivery) to \$9.99 and can be different for different customers as free delivery promotions vary by price area. I observe the delivery fee paid by a customers when she orders online, but I have to impute the delivery fee for trips the agent decided to make in-store. This turns out not to be an issue since the grocer goes "blanket" when it comes to issuing coupons. Therefore, if I observe a household in a particular zipcode having a coupon for discount on delivery in a certain week, I can assume that every other customer living in the same zipcode will have one too. The matrix X_i contains demographic information on the household such as education, employment status and age of the head of the household. *Weekend* is a dummy variable whose purpose is to capture the fact that time pressure is likely to be lower in non working days. μ_i is a random effect that picking up unobservable taste for online trips by agent i. The random effect is distributed as follows

$$\mu_i \sim N(0, \sigma_\mu) \quad (3)$$

Moreover, it is independent from the i.i.d. shock θ_{it} which is a normally distributed disturbance whose variance is normalized to 1. I observe the choice of the shopping channel (where, with c=1, I refer to an online order) which follows the rule

$$c = \begin{cases} 1 & \text{if } z^*_{it} \geq 0 \\ \\ 0 & \text{if } z^*_{it} < 0 \end{cases}$$

4.2 Demand for Cereals

The utility consumer i living in price area p derives from purchasing UPC j in trip t on channel c is modeled as follows for in-store purchases:

$$U_{ijpt} = \beta_0 + \beta_1 Price_{jpt} + \beta_2^{store} * \mathbb{1}\{j \notin h_{it}\} + M_j^{store} + (\delta + \xi_{it}) * \mathbb{1}\{j \notin \Omega_{it}\} + v_{ijt} \quad (4)$$

Similarly, for an online trip:

$$U_{ijpt} = \beta_0 + \beta_1 Price_{jpt} + \beta_2^{online} * \mathbb{1}\{j \notin h_{it}\} + M_j^{online} + (\delta + \xi_{it}) * \mathbb{1}\{j \notin \Omega_{it}\} + v_{ijt} \quad (5)$$

Price is the price per ounce paid by the household, net of discounts; prices for the same items can vary across price areas. The indicator $\mathbb{1}\{j \notin h_{it}\}$ singles out UPCs consumer i never bought in the past. Therefore, β_2 is the impact on utility from purchasing a box of cereals whose UPC did not belong to the shopping history of the household. I allow this effect to be different according to whether the consumer shops online or in a store. M_j are UPC fixed effects, that are also interacted with the channel dummy and pick up the "quality" of a brand. The unobserved shock ξ_{it} only hits brands not experienced by the customer before ($\mathbb{1}\{j \notin \Omega_{it}\}$), representing an instantaneous taste for variety.

The link between the two parts of the model is given by the generating process of the unobserved shock ξ_{it}. I assume that ξ_{it} and θ_{it}, the disturbances in equation (2), are jointly distributed according to a bivariate normal

$$\begin{pmatrix} \theta \\ \xi \end{pmatrix} \sim BN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma\right] \quad \Sigma = \begin{bmatrix} 1 & \rho\sigma_\xi \\ \rho\sigma_\xi & \sigma_\xi^2 \end{bmatrix} \quad (6)$$

The parameter δ can be interpreted as the mean of the ξ_{it} shock. In some specifications, I will parameterize it as follows

$$\delta = \alpha_0 + \alpha_1 Internet + \alpha_2 Weekend \tag{7}$$

This allows the mean taste for new goods to fluctuate according to the channel selected and the time of the week (weekend vs. weekday).

4.3 Discussion

The model captures the most salient features presented in the descriptive section of the paper. The potential bias deriving from sorting of trips across channels is addressed by allowing θ_{it} and ξ_{it} to be correlated. This acknowledges the fact that unobserved determinants of the choice of shopping online (e.g. lack of time) can also affect the attitude towards brand experimentation.

Furthermore, I allow the impact on utility derived from purchase of a particular UPC never bought in the past (β_2) to be different for online and in-store trips. Inertia in the choice of the UPC operates both online and in-store: customers tend to purchase the same box size or variety of their brand of choice over time. For this reason both β_2^{online} and β_2^{store} should be negative. If distaste for UPCs never purchased before is only driven by habits or unobservable characteristics, it should be equal for online and in-store purchases, that is $\beta_2^{online} = \beta_2^{store}$. I interpret the additional stickiness in the choice of the UPC online (that is, the difference between β_2^{online} and β_2^{store}) as the effect of website design, through the shopping history list.

Finally, allowing δ, the mean of the ξ_{it} shock, to vary according to the channel of purchase, I capture structural differences in the taste for new brand for in-store and Internet orders. This difference can be interpreted as representative of higher uncertainty that makes exploration less attractive online.

5 Estimation

Below I describe the estimation procedure for the two components of the model: the selection of the shopping channel and demand for cereals. The large amount of observations at hand, and the importance of unobserved effects in the model make it particularly appropriate to estimate it using Bayesian techniques. In a Gibbs sampling approach, unobserved random effects are drawn from the appropriate distribution rather than integrated out of the likelihood. This allows for significant gains in the estimation time (Train, 2003). I present below sampler for the model including random coefficients; the Gibbs sampler for the model without unobserved heterogeneity is just a simplified version of it. All details are left to the appendix.

5.1 Channel selection: priors and sampling scheme

In order to estimate the channel selection probit in equation (2), I need to specify prior distributions for the parameters of the model: the vector of coefficients γ, and the vector of random effects μ_i. The last parameter of the model is variance of the distribution of the random effects, σ_μ.

$$\sigma_\mu \sim IG(a_1, a_2) \tag{8}$$

$$\mu \sim N(0, \sigma_\mu^0) \tag{9}$$

$$\gamma \sim N(\gamma_0, V_0) \tag{10}$$

a_1, a_2 are known hyperparameters. In particular, the prior on the inverse gamma is chosen to be highly uninformative by setting $a_1 = 1 + 10^{-10}$, $a_2 = 1 + 10^{-5}$.

Sampling from the joint posterior of these parameters directly is challenging. Instead, I apply the Gibbs sampling approach and make draws from the distribution of each parameter at a time, conditioning on the values of the others. This turns out to be much simpler and, after a sufficient number of iterations, the draws from the sequence of the conditional distribution can be assumed to be draws from the joint posterior.

Given initial values for the parameters, the sampling scheme for each iteration is as follows

1. Apply data augmentation (Albert and Chib, 1993) to draw the z_{it}^{*} from a normal truncated distribution.
2. Draw γ from its full conditional distribution, a multivariate normal with mean $\bar{\gamma}$ and covariance matrix V.
3. Draw μ from its full conditional distribution, a normal with mean zero and variance σ_{μ}.
4. Draw σ_{μ} from its full conditional distribution, which is an inverse gamma.

5.2 Demand estimation: priors and sampling scheme

For the demand system, I partition the K-dimensional vector of coefficients into random (β^{i}) and fixed (β') ones. In the application the number of random coefficients is 2; therefore, the dimension of β' is K-2. A random prior is assumed over the distribution of the random coefficients and a prior is imposed over mean and variance of this normal distribution. These priors are, respectively normal and inverse Wishart.

$$\beta^{i} \sim N(b, W) \tag{11}$$

$$b \sim N(m, v) \tag{12}$$

$$W \sim IW(p_1, p_2) \tag{13}$$

The ξ_{it} unobserved shocks are also parameters of the demand model. They are drawn from a bivariate normal jointly with the errors of the channel selection equation as in equation (6). Therefore, the final priors to be specified refer to the mean of the random effect and to the variance covariance matrix of the bivariate normal.

$$\Sigma \sim IW(2, I) \tag{14}$$

where I is the identity matrix.

The last parameter is δ, the mean of the ξ_{it} shocks. In alternative specifications, instead of recovering δ, we will be interested in the coefficients of its parametrization in equation 7, α_0, α_1, and α_2. Once again, I simulate draws from the joint posterior of the parameters by sampling from their conditional distributions. Each iteration t of the sampler unfolds as follows

1. Draw b from a Normal.

2. Draw W from an inverse Wishart.

3. Draw β^i from its conditional distribution

$$\pi(\beta^i|y_{it}, \xi_{it}) \propto L(y|\beta, \xi_{it})\phi(\beta^i|b, W)$$

4. Draw β' from

$$\pi(\beta'|y_{it}, \xi_{it}) \propto L(y_{it}|\beta, \xi_{it})$$

5. Draw ξ_{it} jointly with θ_{it} from the distribution specified in equation (6).

6. Draw Σ from an inverse Wishart.

7. Obtain δ or α's by Bayesian regression.

6 Results

The model is estimated using a random subsample of 500 households[13], for a total of 2,778 trips, 783 of which are online orders. Results are displayed in Table 8 and Table 9 for channel selection and demand respectively.

[13]This is aimed at saving computation time. Given the large number of brands included in my analysis, each shopping situation implies listing of covariates for more than 100 brands. This requires to limit the number of trips taken into consideration to avoid the matrix of regressors from growing too large and slowing down the estimation routine.

Variables enter the channel selection equation with the expected signs. Higher delivery fees make it less appealing to shop on the Internet. Wealth impacts positively on preference for online shopping, whereas the *weekend* dummy is negative and very large (the implied elasticity is 57%). Both results are consistent with the utility of time playing a role in affecting the selection of the shopping channel. As observed by previous literature (Chiou, 2009), customers living further away from a grocery store of the chain have stronger taste for e-commerce. The variables related to age enter with negative coefficient as the reference group age is 18 to 35 years, that is the youngest and most likely to be at ease with technology.

The demand step is estimated under different specifications. In columns 1 and 2 in Table 9 I impose δ, the mean of the shock to the taste for new brands ξ, to be the same for each trip. I relax this assumption in the estimates presented in columns from 3-5, where δ is parameterized as in equation (7). The coefficient on price is negative and close across specifications.

Models reported in columns 2 and 4 include an interaction between the price variable and the channel to assess the existence of different prince sensitivity for online purchase. This interaction enters with positive sign[14] implying that customers are less price sensitive when shopping online[15].

As argued before, the difference in the estimates of the disutility from purchasing a new UPC online and in-store (that is, the difference between β_2^{online} and β_2^{store} in equations (4) and (5)) identifies the effect of web design on brand choice. This magnitude ranges between -.85 and -.91 across specifications. Figure 6 shows the distribution of its draws from the Gibbs sampling:no retained draw was positive and the difference can be as negative as -1.3. The 1% symmetric Bayesian confidence interval for the difference is $[-1.26; -.55]$. The effect of web design can be monetarized using estimates of price sensitivity as follows

$$\text{Dollar value of the list} = \frac{\beta_2^{online} - \beta_2^{store}}{\beta_1} \tag{15}$$

[14] However, on average only about 70% of the draws are positive.

[15] Whereas the early literature has found the opposite (Brynjolfsson and Smith, 2000; Ellison and Ellison, 2009), my result is in line with other studies of the grocery industry (Chu, Chintagunta, and Cebollada, 2008; Andrews and Currim, 2004).

resulting in a dollar value of about 4 dollars. This number may seem large, considering that the average box of cereal in my sample costs \$3.55. It would imply that for a cereal brand outside the *shopping history list* to compete with a brand included in it, it should offer a very large discount (even larger than the price). However, there are two reasons to rationalize the result. First, while I focus on cereal purchases, a shopping trip includes purchases of many other categories; for each of them usage of the list grants the benefit of not having to browse the respective virtual aisle. Therefore, it is hard to think that a discount offered on cereals would be enough to reverse the decision of whether or not to use the list. Second, conditional on the customer using the shopping list, she would not even get to find out about promotions on brands that are not included in her list. This shows clearly that a \$4 discount should, *ceteris paribus*, do nothing to persuade the customer not to use the list.

The parameter δ represents the mean of the unobserved shock to taste for new brands. In columns 1 and 2, the mean is positive, which is not totally surprising since the brand proliferation in the industry is consistent with consumers having taste for variety. However, the size of this effect is tiny compared with the impact of web-design driven state dependence. When I allow the mean of the shock to vary (columns (3) and (4)) it emerges that the mean taste for variety is lower on the Internet, possibly capturing the higher uncertainty over quality faced in online exploration. The parameter α_2 is positive, implying that customers are more likely to value new brands on weekends. This would be the case if, for example, parents are more likely to shop with children during weekends. Finally ρ, the correlation between the unobserved shock to new brands ξ_{it} and the residual from the shopping channel selection, is negative. Large shocks to the instantaneous utility of time (which drive the agent to shop online) are associated with negative shocks to the taste for new brands. This is consistent with the view that online trips are more likely to be made in instances when utility of time is higher.

In making utility derived from consumption of a brand a function of whether or not the agent has bought it before, I am introducing a form of state dependence. It is well known (Heckman, 1981; Dube, Hitsch, and Rossi, 2010) that unobserved heterogeneity can be confounded with state dependence. However, this only bias the estimate of the *level* of state dependence whereas I care about estimating the *difference* between the level state dependence online and in-store. Unless we allow consumers to have different preferences for brands on the two distribution channels (i.e. liking more Rice Krispies online than in-store), estimate of the cross-channel gap in state dependence should not be affected by unobserved heterogeneity. Nevertheless, I estimated the model including random coefficients for β_2^{online} and β_2^{store} (column (5)). As expected, the level of the state dependence changes with respect of the baseline model but the gap between state dependence online and in-store is almost unchanged.

To assess the fit of the model, I compare the Lorenz curve obtained with simulation from my estimates with the one resulting from actual data.[16] Figure 7 shows the fit for brands and for cereal manufacturers, respectively. At the brand level, the model overpredicts concentration on both channels. Still a main stylized fact of the original data is preserved: bigger manufacturers and most popular brands display higher market share for purchases made online. As we would have expected, at the greater level of aggregation the fit improves: the manufacturer level Lorenz curves from actual and fitted data almost perfectly overlap.

7 Counterfactual exercises

In this section, I run some counterfactual experiments in order to understand the practical implications of the estimates of the structural model. In each counterfactual, the initial shopping history of each agent is based on their purchases in the first three months of data (which are therefore not used for the simulation, just as they were not for the estimation). Then, after each trip, the shopping list of the customer is updated to include any eventual new brand she could have picked. The parameters used for the simulations are the ones displayed in column 3 in Table 9. The results I present are averages over 10,000 simulations.

[16] Brynjolfsson, Hu, and Simester (2006) use the Lorenz curve to describe the level of concentration of an industry. In my case, I use the degree at which my simulated Lorenz curve matches the actual one as a synthetic measure of predictive power of the model.

7.1 Shopping cost and the level of brand exploration

Model's estimates pointed to features reducing the cost of shopping online as a major hurdle to online brands exploration. Furthermore, they have by far the highest practical relevance both to the players in the industry and to the policy makers. In fact, their existence depend on a choice about the design of the website which is completely under the control of the retailer, much unlike instantaneous shocks to the utility of time or the intrinsic uncertainty over quality of new items.[17]

How much extra brand exploration would take place on the online channel if the *past shopping history list* feature were not available? Recall that I interpret the difference in the estimate of β_2^{online} and β_2^{store} as the effect of the list. Therefore, simulating a world where the list does not exist implies running the model setting β_2^{online} at the same level as β_2^{store}.[18] This results in the total number of brand trials in online purchases over the two years increasing by 23%. However, in levels, the amount of brand exploration online is still lower than in the store.

7.2 Barriers to entry

Low online brand exploration implies higher entry barriers for new brands on the Internet distribution channel. Consumers seeking to reduce shopping cost will go online to exploit convenient characteristics of the web in that respect, as the "shopping history" list. New entrants are not, by definition, part of such list and will therefore struggle to gain ground. How significant are these barriers to entry?

[17]Lewis (2007) argues that extensive use of text and pictures, as measured by megabytes allocated to the description of an item, can reduce information asymmetries in the case of a buyer unable to personally inspect a car. In this spirit, some design changes could also have the effect of reducing the gap in the uncertainty over new brands between the online and the traditional channel.

[18]Changing the design of the website could affect agents' selection of the shopping channel. This is not captured in the counterfactuals as the selection of the shopping channel is not simulated. Simulating this extra step is unfeasible as nothing in the data identifies the impact of changes in design on taste for the online channel: the grocer's website maintained the same structure all throughout the sample period. The changes considered in the counterfactuals would probably reduce the appeal of the Internet service because it limits its potential to reduce the cost of shopping. Therefore, the reported increases in the amount of online exploration can be seen as upper bounds.

To assess this I modified the initial shopping histories of all consumers so that Cinnamon Toast Crunch (a fairly popular cereal) is in nobody's "shopping history" list. This amounts to simulating the entry of a new brand looking exactly as Cinnamon Toast Crunch (same price and characteristics) but lacking any installed base.[19] Therefore, results from this counterfactual can also be interpreted as an evaluation of the difference in entry barriers faced by a new entrant online and in-store.

Figure 8 compares Cinnamon Toast Crunch's market shares under the baseline model and the counterfactual on both distribution channels. At first, the counterfactual market share is lower than in the baseline, both online and in-store. This is expected: the brands starts without any installed base of customers. The figures tend then to converge over time; however this catch-up is faster in-store. Figure 9 displays the ratio between the counterfactual and the baseline market share over time on the two channels. The new brand is performing similarly over the two channel in the first semester. However, Cinnamon Toast Crunch eventually recovers as a new entrant 60% of its in-store market share in the baseline case whereas it never exceeds 40% of its online one. The only difference between the two environments is that a new brand, on top of any entry barrier it faces in-store (quantified by the level of β_2^{store} in my estimates) faces the additional effect of the list (the difference between β_2^{online} and β_2^{store}) on the Internet channel. The result is even more striking since the wedge caused by the list manages to slow down adoption of a brand that was fairly well liked by consumers to begin with.

7.3 Contextual advertising/Recommendations

Contextual advertising and customer recommendations are popular form of advertising and customer loyalty enhancing for online businesses. Context ads are individually targeted promotion based on the content of the page the consumer is browsing. For example, they are regularly displayed along with results from a search on a search engine. Recommendation systems offer suggestions to a customer based on her previous shopping history. The firm

[19] To be precise, the exercise simulates the entry of such a brand in a world where Cinnamon Toast Crunch did not exist.

uses purchases by other customers with similar shopping history to forecast goods or offers that may be of interest to her. This system has been made popular by Amazon.com and Netflix. Inclusion of recommendations or context ads in the design of the website can make it easy for consumers to notice brands they have not tried before and, therefore, encourage exploration.

I simulate a scenario in which, while looking at their "shopping history" list, consumers are offered the chance of purchasing a cereal brand that is not part of it. Figure 10 provides a graphical representation of this situation. In the same screen with the list of already purchased brand, the customer can observe (on the right) another brand she has not purchased before. To implement this exercise I once again remove Cinnamon Toast Crunch from the brand history of every household (making it look like a new entrant). However, I set β_2^{online} at the same level as β_2^{store} only for the Cinnamon Toast Crunch brand. This means that the effort required to a household in order to purchase Cinnamon Toast Crunch, is the same as the one needed to purchase another brand already part of the household's shopping history list. This would be the case if the brand were feature in a recommendation box such as the one portrayed in Figure 10.

Figure 11 displays the ratio between market share in the counterfactual and in the baseline simulation over time for the two channels. The convergence to the benchmark market share evolves pretty much like in the previous experiment for in-store sales; there is a marked difference for the results on the online channel. Online market share for the newly introduced brand reaches the same level (actually it exceeds it) as in the baseline model in the first semester. Therefore, with the help of a recommendation system, it would take less than six months for Cinnamon Toast Crunch to make up for the disadvantage of being stripped of its initial share of loyal customers online. Moreover, the online market share in the counterfactual stabilizes at a level 50% higher than the one held in the baseline on the same channel. This suggests that not only the context ad helps the brand to regain its customer base but also attracts new customers.

8 Conclusions

In this paper, I focus on how the choice of the shopping channel affects households' propensity to try brands they have not purchased in the past. After documenting that brand exploration is more prominent in-store than online, I investigate the role of three different mechanisms that can explain this result. My estimates suggest that popular web features making it more quick and convenient to shop online (such as "shopping history" lists) play a relevant role in explaining why brand trial is less pervasive on the Internet.

In the counterfactuals, I assess the degree to which this situation implicitly creates higher barrier to entry online by simulating the introduction of a new cereal brand. I find that penetration of the market is significantly harder online for a new entrant. However, I also show that a different design of the website (i.e. the the introduction of context ads and recommendation system) has the potential of reversing this effect, making it easier for the "outsiders" to become popular online.

Conventional wisdom has it that the Internet removes traditional sources of switching cost and monopoly power (i.e. location advantage), making competition fiercer and lowering barriers to entry. However, I show an instance in which e-commerce has the opposite effect, introducing new barriers of its own. This results from the presence of a "shopping history" list on the grocer website; however, identical or similar features can be found on a large number of e-commerce platforms. Therefore, my findings apply to wide variety of online shopping contexts. In this respect, my work ties into a growing body of literature deepening the analysis of the Internet medium past the initial characterization as a shock reducing search costs.

References

Ackerberg, D., K. Hirano, and Q. Shahriar (2009): "The Buy-it-now Option, Risk Aversion and Impatience in an Empirical Model of eBay Bidding," working paper.

Aguiar, M., and E. Hurst (2007): "Life-Cycle Prices and Production," *American Economic Review*, 97(5), 1533–1559.

Albert, J. H., and S. Chib (1993): "Bayesian analysis of binary and polychotomous response data.," *Journal of the American Statistical Association*, 88(422), 669–679.

Andrews, R. L., and I. S. Currim (2004): "Behavioral Differences between Consumers Attracted to Shopping Online vs. Traditional Supermarkets: Implications for Enterprise Design and Strategy," *International Journal of Marketing and Advertising*, 1(1), 38–61.

Bakos, J. Y. (1997): "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science*, 43(12), 1676–1692.

Bayus, B. L. (1992): "Brand loyalty and marketing Strategy: An application to home appliances," *Marketing Science*, 11, 21–38.

Bhatnagar, A., S. Misra, and H. R. Rao (2000): "On Risk, Convenience, and Internet Shopping Behavior.," *Communications of the ACM*, 43(11), 98–105.

Broda, C., and D. Weinstein (2009): "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, forthcoming.

Brown, J. R., and A. Goolsbee (2002): "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *The Journal of Political Economy*, 110(3), 481–507.

Brynjolfsson, E., A. A. Dick, and M. D. Smith (2010): "A Nearly Perfect Market?," *Quantitative Marketing and Economics*, 8(1), 1–33.

Brynjolfsson, E., Y. J. Hu, and D. Simester (2006): “Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on Concentration of Product Sales,” MIT Center for Digital Business Working Paper.

Brynjolfsson, E., and M. D. Smith (2000): “Frictionless Commerce? A Comparison of Internet and Conventional Retailers.,” *Management Science*, 46(4), 563–585.

Burke, R. R., B. A. Harlam, B. E. Kahn, and L. M. Lodish (1992): “Comparing Dynamic Consumer Choice in Real and Computer-simulated Environments.,” *Journal of Consumer Research*, 19(1), 71–82.

Chen, P.-Y., and L. M. Hitt (2002): “Measuring Switching Costs and the Determinants of Customer Retention in Internet-Enabled Businesses: A Study of the Online Brokerage Industry.,” *Information Systems Research*, 13(3), 255–274.

Chintagunta, P., J. Chu, and J. Cebollada (2009): “What Drives Channel Choice in Grocery Shopping?,” working paper.

Chiou, L. (2009): “Empirical Analysis of Competition between Wal-Mart and Other Retail Channels,” *Journal of Economics and Management Strategy*, 18, 285–322.

Chu, J., P. Chintagunta, and J. Cebollada (2008): “A Comparison of Within-Household Price Sensitivity across Online and Offline Channels,” *Marketing Science*, 27, 283–299.

Clemons, E. K., I.-H. Hann, and L. M. Hitt (2002): “Price Dispersion and Differentiation in Online Travel: An Empirical Investigation.,” *Management Science*, 48(4), 534–549.

Crawford, G. S., and M. Shum (2005): “Uncertainty and Learning in Pharmaceutical Demand,” *Econometrica*, 73(4), 1137–1173.

Danaher, P. J., I. W. Wilson, and R. A. Davis (2003): “A Comparison of Online and Offline Consumer Brand Loyalty.,” *Marketing Science*, 22(4), 461–476.

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(1), 55–78.

Dube, J. P., G. Hitsch, and P. Rossi (2009): "Do Switching Costs Make Markets Less Competitive?," *Journal of Marketing Research*, 46, 1–22.

——— (2010): "State Dependence and Alternative Explanations for Consumer Inertia," *RAND Journal of Economics*, forthcoming.

Einav, L., E. Leibtag, and A. Nevo (2010): "Recording Discrepancies in Nielsen Homescan Data: Are They Present and Do They Matter?," *Quantitative Marketing and Economics*, 8(2), 207–239.

Ellison, G., and S. Ellison (2009): "Search, Obfuscation, and Price Elasticities on the Internet," *Econometrica*, 77, 427–452.

Erdem, T., and M. Keane (1996): "Decision Making under Uncertainty: Capturing Dynamic Choice Process in Turbulent Consumer Goods Markets," *Marketing Science*, 15(1), 1–20.

Goettler, R. L., and K. Clay (2006): "Tariff Choice with Consumer Learning:Sorting-Induced Biases and Illusive Surplus," working paper.

Hann, I.-H., and C. Terwiesch (2003): "Measuring the Frictional Costs of Online Transactions: The Case of a Name-Your-Own-Price Channel.," *Management Science*, 49(11), 1563–1579.

Heckman, J. (1981): *The Incidental Parameter Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process and Some Monte Carlo Evidence*in "Structural Analysis of Discrete Data with Econometric Applications", Manski, C. and McFadden, D. (eds), MIT Press.

Hong, H., and M. Shum (2006): "Using price distributions to estimate search costs," *RAND Journal of Economics*, 37, 257–275.

Jin, G., and E. Kato (2007): "Dividing Online and Offline: A Case Study," *Review of Economics Studies*, 74, 981–1004.

Johnson, E. J., S. Bellman, and G. Lohse (2003): "Cognitive Lock-In and the Power Law of Practice," *Journal of Marketing*, 67, 62–75.

Klemperer, P. (1995): "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economics Studies*, 62, 515–539.

Lewis, G. (2007): "Asymmetric Information, Adverse Selection and Seller Disclosure: The Case of eBay Motors," working paper.

Mazar, N., A. Herrmann, and E. J. Johnson (2007): "Preferences and choice behavior online: A cognitive cost approach to understanding product class differences," working paper.

Nevo, A., and C. Wolfram (2002): "Why Do Manufacturers Issue Coupons? An Empirical Analysis of Breakfast Cereals," *RAND Journal of Economics*, 33(2), 319–339.

Pozzi, A. (2009): "Right to your Door: Home Delivery Service and Customers Price Sensitivity for Heavy Items," working paper.

Scherer, F. M. (1979): "The Welfare Economics of Product Variety: An Application to the ready-to-eat Cereals Industry.," *Journal of Industrial Economics*, 28(2), 113–134.

Schmalensee, R. (1974): "Brand Loyalty and Barriers to Entry," *Southern Economics Journal*, 40, 579–588.

——— (1978): "Entry deterrence in the ready-to-eat breakfast cereal industry.," *Bell Journal of Economics*, 9(2), 305–327.

Shin, S., S. Misra, and D. Horsky (2007): "Disentangling Preferences, Inertia and Learning in Brand Choice Models," working paper.

Shum, M. (2004): "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereal Market," *Journal of Economics and Management Strategy*, 13(2), 241–272.

Train, K. (2003): *Discrete Choice Methods with Simulation*. Cambridge University Press.

Tables and Figures

Table 1: Trip descriptive statistics by channel of purchase, for all the trips.

	All trips			**Large trips**		
	In store	*Online*	*P-value*	*In store*	*Online*	*P-value*
Total expenditure	45.31	151.99	.000	162.16	179.70	.000
Basket size	15.15	55.61	.000	69.59	77.85	.000
Number of unique items	11.35	34.34	.000	62.41	63.05	.000

Total expenditure is net of any discount and is expressed in dollars. The total number of in-store trips is 1,662,375; while the total number of online orders is 166,879. Basket size refers to the total number of items purchased in the trip; as opposed to the number of unique item, that does not double count multiple purchases of the exact same item. "Large trips" are defined as trips worth more than 100 dollars (212,946 in store trips and 122,320 online), when comparing total expenditure; or trips with a basket size of at least 50, when comparing basket size (116,236 in store trips and 84,148 online); or trips at least 50 unique items purchased, when comparing number of unique items (49,061 in store trips and 28,196 online). The P-values refer to a t-test for the equality of means.

Table 2: Product categories most frequently purchased in in store and online trips.

In store	*Online*
Milk and substitutes	Milk and substitutes
Carbonated soft drinks	Fresh bread
Fresh bread	Bananas
Bananas	Carbonated soft drinks
Refrigerated yogurt	Salad vegetables
Salad vegetables	**Cold cereals**
Cold cereals	Eggs and substitutes
Cooking vegetables	Refrigerated yogurt
Still water	Still water

The ranking is computed by counting the number of trips in which *at least one* item of a given product category has been purchased.

Table 3: Trip descriptive statistics by channel of purchase, for all the trips involving purchase of breakfast cereals.

	In store	*Online*
Total expenditure	110.83	179.24
	(78.61)	(89.53)
Basket size	40.51	66.29
	(27.77)	(35.12)
Number of unique items	30.56	41.56
	(20.03)	(19.04)
Number of unique categories	23.29	30.42
	(14.36)	(12.65)
Number of cereal boxes	1.52	1.88
	(.8913)	(1.49)
Number of unique cereal brands	1.26	1.40
	(.5640)	(.8043)
Number of obs.	106,378	35,647

Total expenditure is net of any discount and is expressed in dollars. Basket size refers to the total number of items purchased in the trip; as opposed to the number of unique item that does not double count multiple purchases of the exact same item. Standard deviations are reported in parentheses.

Table 4: Demographic information.

	5^{th} *percentile*	25^{th} *percentile*	50^{th} *percentile*	75^{th} *percentile*	95^{th} *percentile*
Number of trips	43.8	96	158	245	449.2
Number of cereal trips	3	10	21	36	68
Share trips online (%)	.4	1.7	5.6	16.7	53.6
Share cereal trips online (%)	0	1.4	12.5	40	.91
Black	0	.2	1.7	4.5	19.7
Hispanic	.4	3.8	7.4	14.4	36.6
College degree	18.8	35	49.6	63.2	78.7
Employed	49.7	60.4	66.7	71.9	79.2
Per capita income	16,261	23,837	30,791	41,884	69,045
Commute $\leq$ 30 min	36.18	49.3	57.3	66	79.3
Commute 30 to 59 min	14.7	25.1	32.2	38.6	49.3
Commute $>$ 60 min	1.5	5	8.9	14.3	23.7
$15 <$ Age < 35	0	4.6	9.7	17	31.6
$35 \leq$ Age < 54	27.1	44.5	53.6	62.5	75.1
$54 \leq$ Age < 65	0	10.4	16.1	22.1	32.9
Age ≥ 65	0	8.2	15.8	25	42.2
Distance	.26	.65	1.09	1.72	3.38

Information on number of trip, number of cereal trips and number of online trips are constructed from the scan data provided from the grocer. Distance is is calculated as miles between the domicile of the household and the closest brick-and-mortar store of the supermarket chain and was also provided by the retailer. All the other variables (*black, hispanic, college degree, employed, per capita income, age and commuting time*) are matched from Census 2000 at the block group level for the 6,021 households for which 9-digit zipcode of residency was provided.

Table 5: Brand and UPC exploration for in-store vs. online trips.

Panel A: Brand exploration

	In store	*Online*	P-value
All trips	.3576	.2593	.000
	(.0013)	(.0019)	
Trips worth > $100	.3517	.2600	.000
	(.0018)	(.0020)	
Basket size > 50	.3568	.2652	.000
	(.0030)	(.0034)	

Panel B: UPC exploration

	In store	*Online*
Same brand, same UPC	56,296	29,630
(num of trips)	(65.1%)	(79.66%)
Same brand, different UPC	30,177	7,563
(num of trips)	(34.9%)	(20.34%)

Brand trial is defined as purchase of a brand not bought in the previous three months; UPC exploration is analogously defined. Panel B conditions on trips were the household did not engage in brand exploration. P-values refer to t-test where the null hypothesis is equality of the means. Standard errors are reported in parenthesis.

Table 6: Probability of exploration.

Variable	**(1)** Dep.Var. New brand	**(2)** Dep.Var. New brand	**(3)** Dep.Var. New brand	**(4)** Dep.Var. New brand	**(5)** Dep.Var. New upc	**(6)** Dep.Var. New upc
online	-.129**	-.115**	-.117**	-.077**	-.099**	-.110**
	(.0066)	(.0069)	(.0069)	(.0035)	(.0050)	(.0036)
black		.0001	.0002		.0004	
		(.0004)	(.0004)		(.0004)	
hispanic		-.0005	-.0004		.0001	
		(.0003)	(.0003)		(.0004)	
family		-.0006*	-.0006*		-.0009**	
		(.0003)	(.0003)		(.0003)	
college degree		-.0013**	-.0014**		-.0010**	
		(.0003)	(.0003)		(.0002)	
income pc		-.0001	-.0001		-.0001	
		(.0003)	(.0003)		(.0003)	
age35_54		-.0003	-.0003*		-.0004	
		(.0003)	(.0003)		(.0003)	
age55_64		-.0003	-.0003		-.0003	
		(.0004)	(.0004)		(.0003)	
age≥65		-.0010*	-.0010**		-.0006	
		(.0004)	(.0004)		(.0004)	
distance		.0044	.0044			
		(.0036)	(.0035)			
day of the week f.e.	No	No	Yes	Yes	Yes	Yes
household f.e.	No	No	No	Yes	No	Yes
observations	163,324	117,062	117,062	163,324	102,108	129,333

Trial of a new brand is defined as purchase of a brand not bought in the previous 3 months; trial of a new upc is analogously defined. Regressions for new upc's trial only include trips were the household did not try a new brand. Columns 1,2,3,5 are probit and marginal effects are reported. Column 4 and 6 are linear probability models. Standard errors in parentheses. Significance levels :$*$: 5% $**$: 1%

Table 7: Robustness checks.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
online	-.081**	-.042**	-.084**	-.068**	-.132**	-.063**	-.116**
	(.007)	(.003)	(.005)	(.004)	(.007)	(.005)	(.034)
trip size		.000					
	(.001)						
promotion			.024**				
			(.003)				
promotion× online			.014**				
			(.005)				
family × online							.000
							(.001)
Household f.e.	X	X	X	X		X	
Brand f.e.				X			
Brand×Week f.e.					X		
Obs.	29,898	163,324	163,324	163,324	121,148	114,926	117,062
Sample	>50 trips	all	all	all	top 25 brands	no kid cereals	demo

Trial of a new brand is defined as purchase of a brand not bought in the previous 3 months. Trips size is measured as the worth of the trip in dollars. Standard errors in parentheses. Column 1 displays estimates for the subsample of households shopping for cereals at least 50 times. Column 6 excludes all trips involving purchase of kids'cereals. Column 7 only includes households for which demographic information is available. Columns 1-6 are linear probability models, Column 7 is a probit with marginal effects reported. Standard errors clustered at the household level.Significance levels :$*$: 5% $**$: 1%

Table 8: Channel selection equation estimates.

Variable	Coef.	Variable	Coef.
items count	.0002	income per cap	.0049
	(.0002)		(.0004)
distance	.0087	age35_54	-.0030
	(.0009)		(.0002)
fee	-.1001	age55_64	-.0071
	(.0006)		(.0003)
weekend	-.4929	age≥65	-.0028
	(.0002)		(.0002)
married	-.0086	college	-.0036
	(.0001)		(.0045)
employed	-.0003	σ_μ	.124
	(.0003)		(.0011)

The unit of observation is a trip (2,778 trips). Burn-in period: 10,000 draws, mean and standard deviation of the posterior are computed on the basis of the next 10,000 draws. Marginal effects reported.

Table 9: Demand estimates.

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
price	-.21	-.22	-.21	-.22	-.15
	(.0206)	(.0248)	(.0213)	(.0239)	(.0272)
price*Internet		.09		.09	
		(.0488)		(.0485)	
$\mathbb{I}\{j \notin h_{it}\} * Online_t \quad (\beta_2^{online})$	-6.96	-7.01	-6.88	-6.88	-4.33
	(.1247)	(.1340)	(.1283)	(.1239)	(.1113)
$\mathbb{I}\{j \notin h_{it}\} * (1 - Online_t) \quad (\beta_2^{store})$	-6.11	-6.11	-5.94	-5.95	-3.35
	(.0661)	(.0695)	(.0668)	(.0701)	(.1032)
δ	.93	.94			
	(.0487)	(.0474)			
α_0			.54	.54	.61
			(.0382)	(.0391)	(.0371)
α_1			-.70	-.70	-.47
			(.1744)	(.1732)	(.1657)
α_2			.07	.06	.08
			(.0378)	(.0299)	(.0383)
σ_ξ	1.13	1.16	1.58	1.35	2.11
	(.0585)	(.0588)	(.0609)	(.0601)	(.3392)
ρ	-.13	-.13	-.16	-.16	-.26
	(.0553)	(.0579)	(.0576)	(.0588)	(.0578)

The unit of observation is a cereal trip (2,778 observations). Standard deviation of the posterior is reported in parenthesis. Burn-in period: 10,000 draws, mean and standard deviation of the posterior are computed on the basis of 10,000 draws of which 1 out of 10 is retained. UPC fixed effects are included in the specification and interacted with the channel dummy. ρ is the correlation between the residuals of the channel selection equation and the unobserved shock to utility derived from new brands. δ is the mean of the unobserved shock to utility derived from new brands, which is estimated as a parameter in column 1 and 2. In columns 3-5, δ is assumed to be a function of the channel and of the time of the trip and the coefficients of this parametrization α_0, α_1, α_2 are reported.

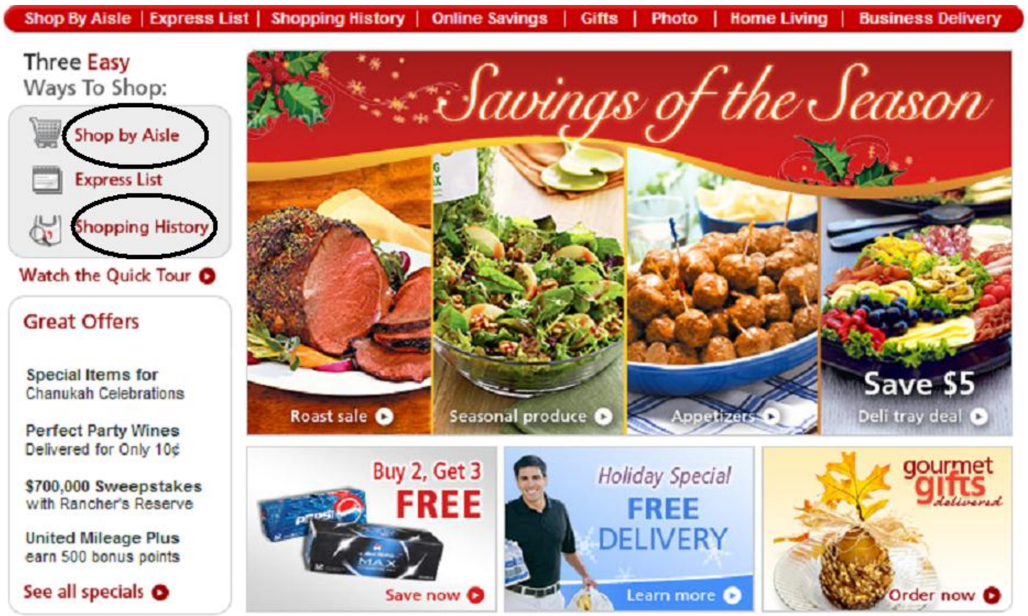


Figure 1: Screenshot from the grocer 's website.

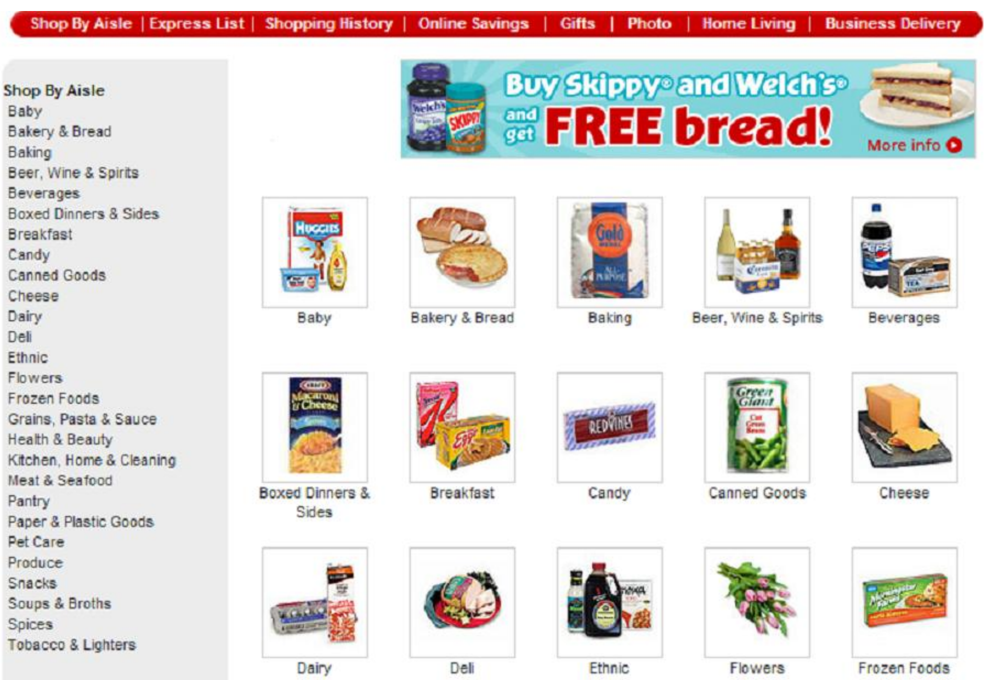


Figure 2: Screenshot from the grocer 's website: shopping by aisle.

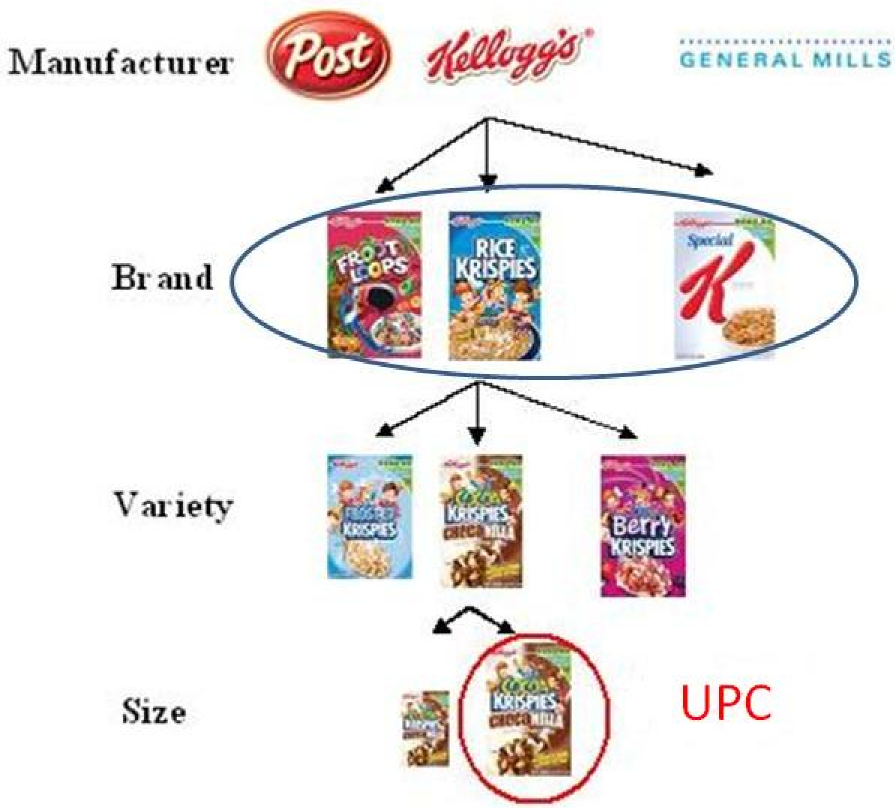


Figure 3: Structure of the data: brand vs. UPC.

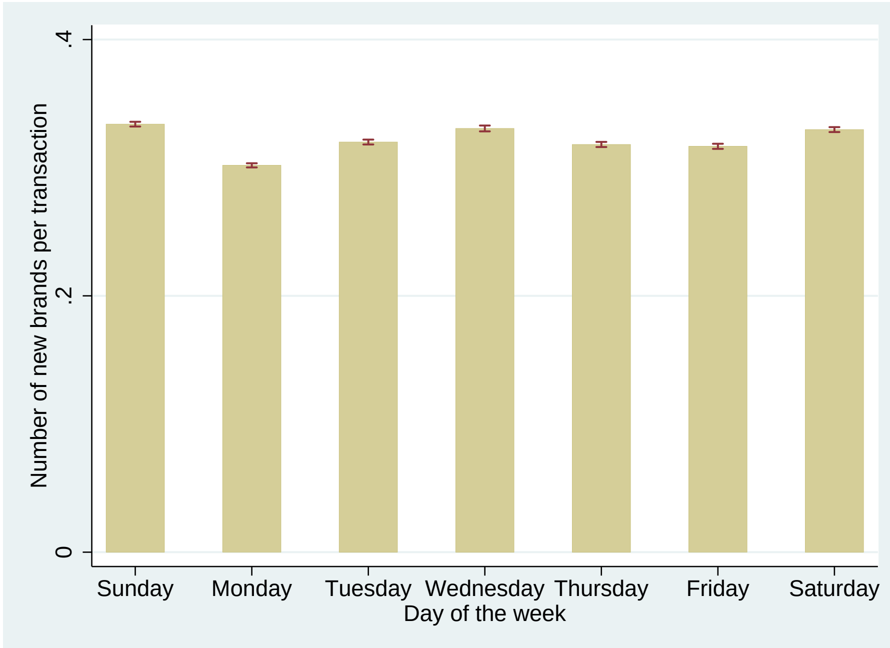


Figure 4: Number of new trials per trip, by day of the week. Confidence intervals are displayed on top of each bar.

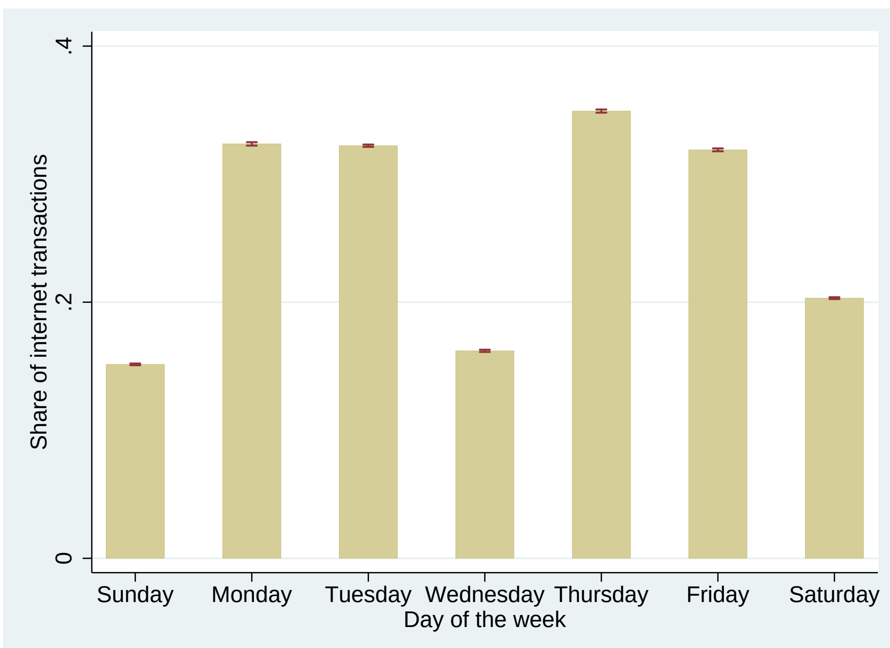


Figure 5: Share of Internet trips, by day of the week (delivery day). Confidence intervals are displayed on top of each bar.

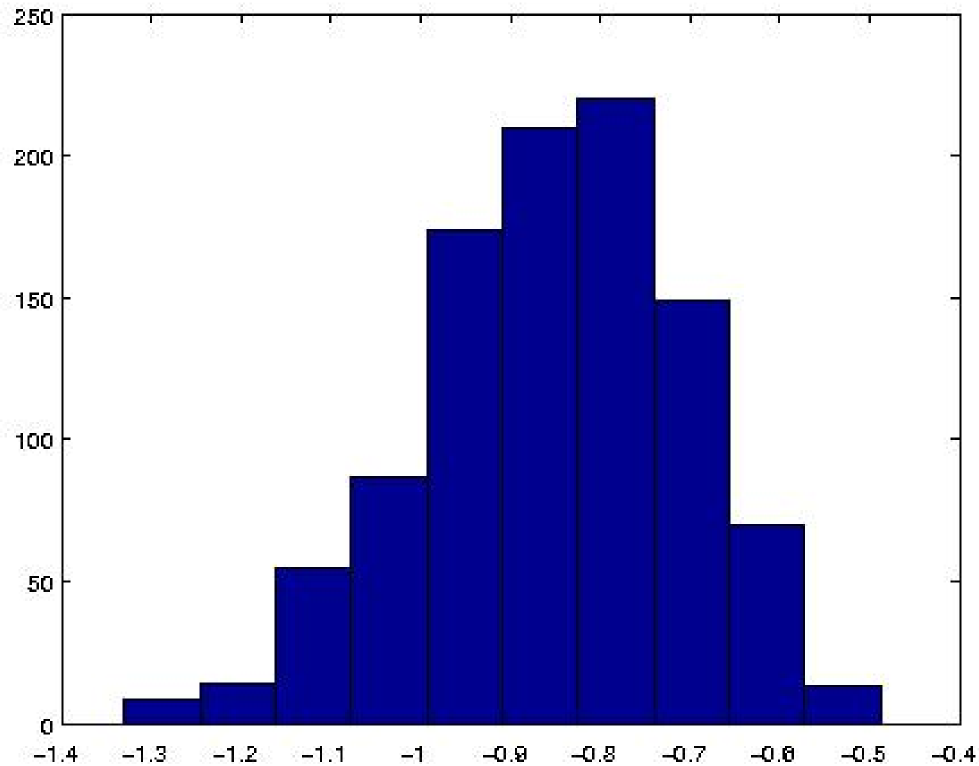
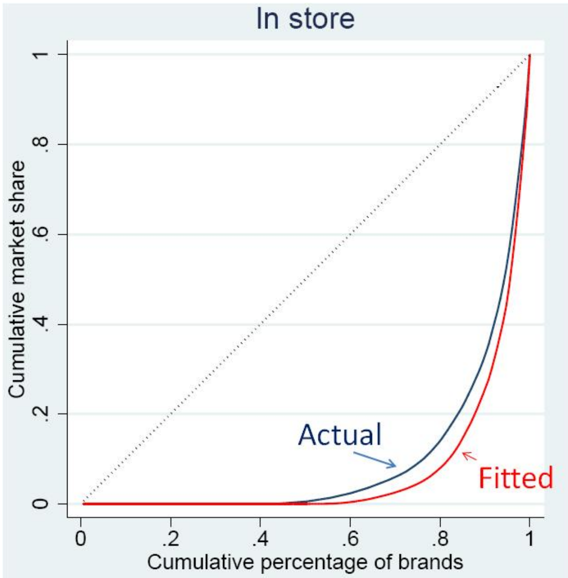
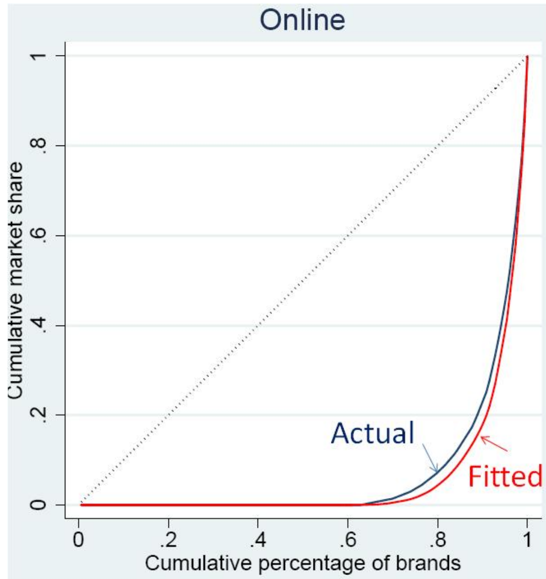


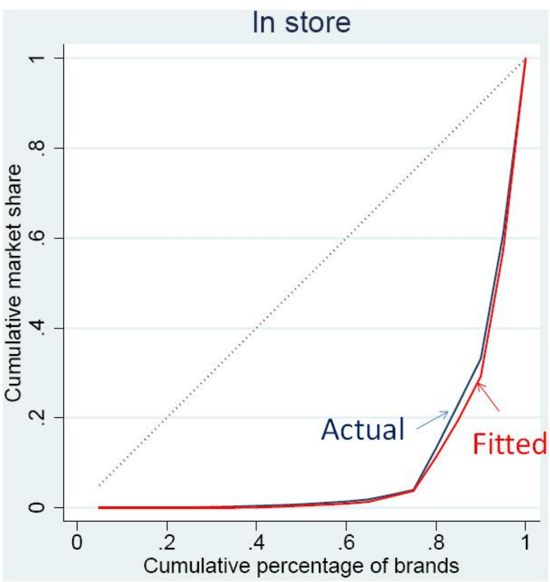
Figure 6: Distribution of the effect of the shopping history list, measured as the difference between β_2^{online} and β_2^{store}. The figure refers to the results from column 1 of Table 9. The number of retained draws is 2,500.



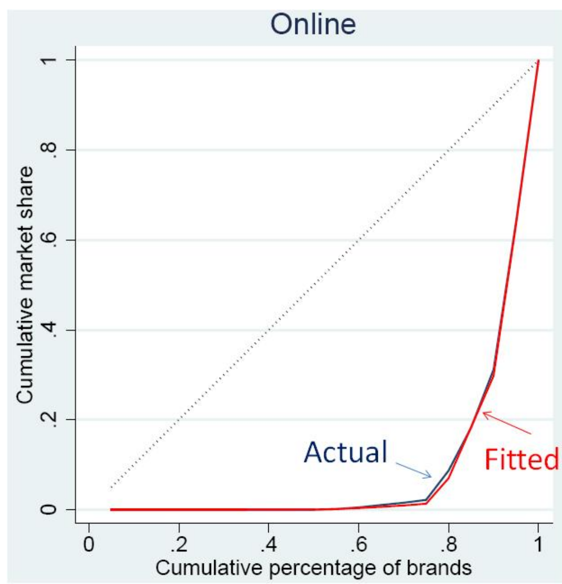
(a) Lorenz curve for brands, in-store sales



(b) Lorenz curve for brands, online sales



(c) Lorenz curve for manufacturers, in-store sales



(d) Lorenz curve for manufacturers, online sales

Figure 7: Fit of the model

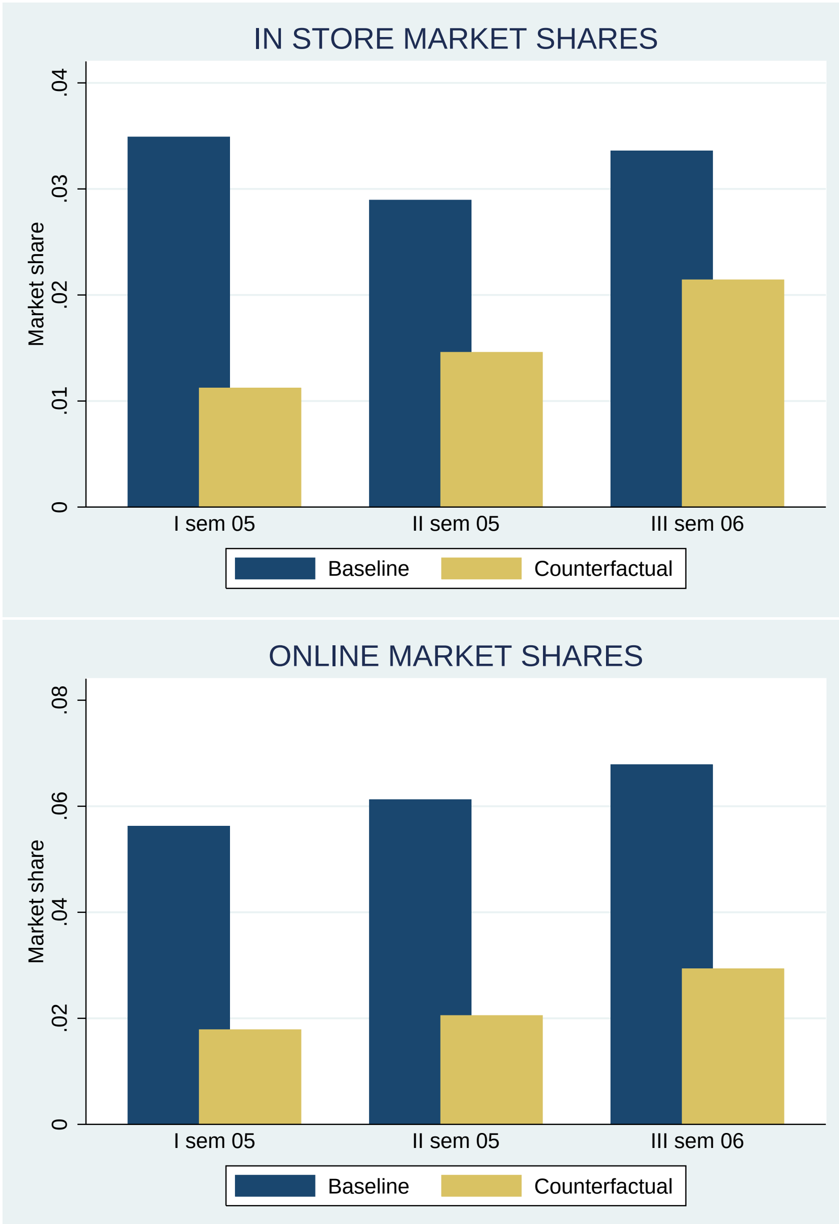


Figure 8: The top panel compares in-store market share for Cinnamon Toast Crunch in the baseline simulation and in the counterfactual where it is removed from everybody‘s shopping history list (de facto treated as a new brand). The bottom one presents the same comparison for online sales.

Figure 9: The bars represent Cinnamon Toast Crunch's market share in the counterfactual as a percentage of the same figure in the baseline model.

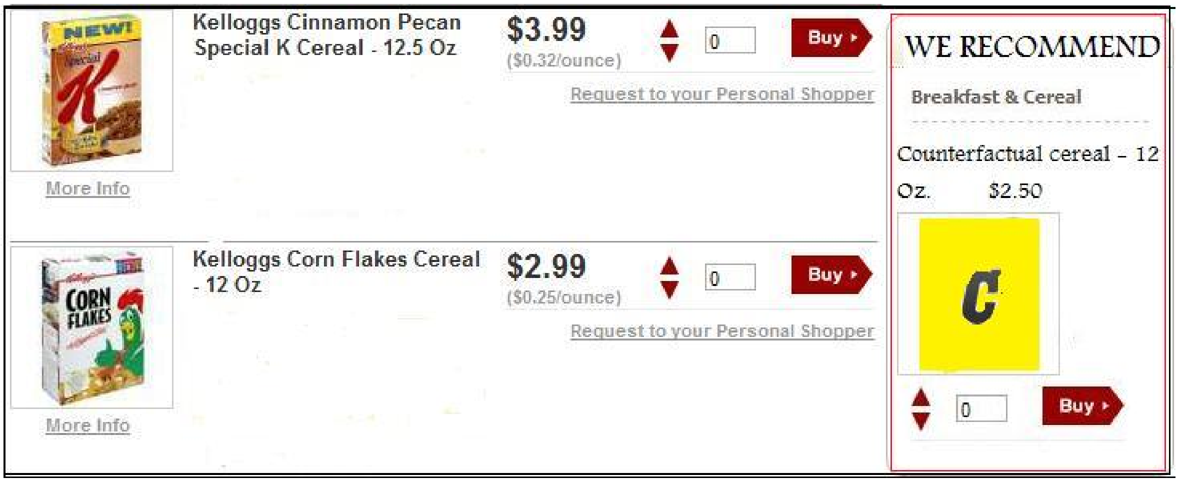


Figure 10: Screenshot of the shopping history list page, featuring context ads.

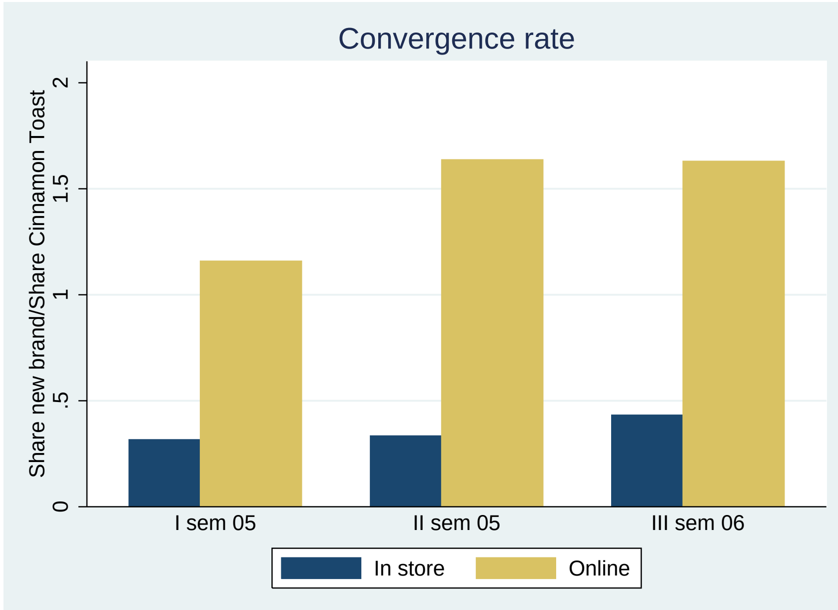


Figure 11: The bars represent Cinnamon Toast Crunch's market share in the counterfactual with context ads as a percentage of the same figure in the baseline model.

Appendix - Not for publication

A-1 Shopping history and initial conditions

Precisely identifying brand exploration using transaction data is problematic because of left-truncation: we do not observe the brands the household purchased before the beginning of the data series. To solve this problem I adopt a simple approach to simulate the set of brands known to the consumer at time t_0. I use the first three months of data to calibrate such set and exploit only the remaining 21 months for the estimation. This is somewhat arbitrary and subject to the usual problems faced in defining initial conditions (Erdem and Keane, 1996).

The choice of calibrating initial conditions on a three months window results from a trade-off. On one hand, the longer the span of data allocated to recovering the set of known brands the more accurate it will be. On the other hand, extending the length of the calibration window reduces the amount of data left for the estimation. The assumption implicit in the choice is that three months is a long enough period to credibly recover the set of cereal brands explored by the household in the past. How many brand exploration instances are spuriously generated because of this assumption? How would things change if I had a longer panel to calibrate the initial conditions? Here I offer two distinct pieces of evidence in support of my results.

For the first exercise, I rely on a scanner dataset different from the one used in the main analysis. The data come from the HomeScan panel, collected by AC Nielsen. Each household in the sample is equipped at home with a scanning device through which they can record all their purchases at every grocery chain visited.[20] The sample I obtained covers over 89,000 households in the period between January 1998 and December 2007, although not all the households are in the sample for the entire span. 85,827 households shopped for breakfast cereals at least once for a total of 2,752,802 trips. The number of households engaging in online shopping of grocery is limited, which explains why this data source could not be used for the core analysis. The availability of a much longer panel partially relaxes the constraint that forced me to use a relatively short period to calibrate the initial conditions. I restrict my attention to the time period between June 2004 and June 2006, the same time span covered by the main data. Using the same definition of brand exploration adopted so far, I can calculate the number of brand trials performed by households in the HomeScan sample under two different settings. The first one calibrates the set of already known brands using the June 2004 - September 2004 purchases, just as in the rest of the paper. Then, I compute the amount of exploration using the entire shopping history prior to September 2004 to set the initial conditions. For some households this amounts to having as many as six years of purchases to identify the set of previously purchased brands.

Table A-1 compares the results. For the purchases generated by households active between June 2004 and June 2006, the difference between the number of brand trials counted using a three months calibration span and the number resulting from using the full previous history for the calibration is 71,194 cases. This implies that mislabeled brand exploration instances make up 3% of the overall number of trips. Some of these households, however, may not have been in the sample very long before June 2004 which would reduce the gap between using three months or the entire history to define brand exploration.

A more powerful test implies computing the same statistics for a subsample of households who have a fairly long record of past purchases prior to September 2004. The second column of the table conditions to household that have been in the sample for at least four years as of June 2006; implying that their shopping history can be recovered from at least two years of past purchases. This increases the estimated inaccuracy of the initial shopping history based on a three months interval. Now the misclassified brand trial amount to 10% of the total number of transactions. This figure maintains stable if we strengthen the requirement focusing on households who have been the sample for at least 6 or 8 years.

It is hard to judge whether 10% is a high or low number and it is unclear to what extent the results obtained based on the HomeScan sample can be carried over to the main piece of data used for the estimation of the model. Nevertheless, this simple exercise suggests that the amount of spurious brand exploration generated by lack of a long panel is not dramatically large.

[20] See Aguiar and Hurst (2007); Broda and Weinstein (2009); Einav, Leibtag, and Nevo (2010) for more detailed descriptions of this data source.

Table A-1: Divergences in the number of new trials for different length of window used to calibrate the initial conditions.

Length of the period used for the calibration	*Full sample*	*Households in the sample for at least 4 years*	*Households in the sample for at least 6 years*	*Households in the sample for at least 8 years*
Number of divergences	71,194	26,394	19,174	10,946
Percentage of divergences	3%	10%	10%	10%
Number of households	47,250	5,827	3,984	1,853

A divergence is defined as an instance where purchase of a brand is labeled as an exploration event relying on a three-months window to calibrate the set of brands known to the household, whereas it would not qualify as such if the set had been constructed relying on the entire prior shopping history available.

The second exercise exploits the grocer's data and checks the robustness of the results presented in Table 6 to changing the length of the interval used to retrieve the set of initially known brands. Being limited by the overall length of the panel, I cannot vary the calibration interval as widely as when using HoeScan data. I estimate the model in equation 1 calibrating initial conditions on intervals of length 1month, 3 months (as in the paper), 6 months, and 1 year. The specifications include household and day-of-the-week fixed effects.

In Figure A-1 I plot the resulting coefficients. To allow for comparison between results emerging from different interval lengths, the coefficients are normalized by the average of the dependent variable. Results are similar with online purchases determining a reduction in the amount of brand exploration between 20% and 30%, with confidence intervals overlapping. The one exception is for results based on previous history calibrated over a single month of purchases: the effect there is lower. However, the fact that the average interpurchase time for breakfast cereal is over two weeks implies that in that specification we are tried to infer which cereals the consumer knows already based on a couple of purchases, which makes it an extreme benchmark.

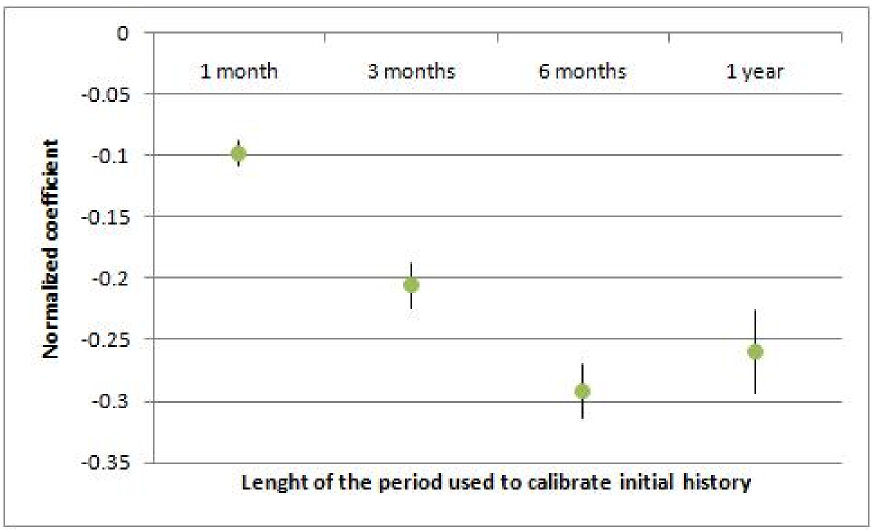


Figure A-1: Coefficient on the *Online* dummy from estimation of equation 1, for construction of the previously known brands exploiting different length of previous shopping history. The picture compares results obtained using one, three, six months or one year of purchases. The regression coefficients are normalized by the mean of the dependent variable to allow for comparison.

A-2 Estimation

Details related to the estimation routine presented in Section 5 are discussed below.

A-2.1 Selection of the Shopping Channel

Draw of the z^*_{it}

The distributional assumption made over the θ_{it} in equation (6), implies that the z^*_{it} are distributed according to a truncated normal with mean $C_{it}\gamma + \mu_i$ and variance 1. The support is $(-\infty, 0]$ if the trip occurs in a store, and $[0, +\infty)$ if the trip is made online.

Draw of γ

The full conditional posterior on γ is normal with mean $\bar{\gamma}$ and variance matrix V. Given the analogy with the posterior for regression parameters in a linear model, we have

$$\gamma \sim N((C'C)^{-1}C'Z^*, (C'C)^{-1}) \tag{A-1}$$

where Z^* is the vector of the stacked z^*_{it}.

Draw of μ

For each agent, the random effect μ_i is drawn from a normal distribution with mean 0 and variance σ_μ.

Draw of σ_μ

The density of σ_μ conditional on the other parameters is an inverse gamma distribution, $IG(\hat{a_1}, \hat{a_2})$, where

$$\hat{a_1} = \frac{N_I + 2a_1}{2} \tag{A-2}$$

$$\hat{a_2} = \frac{\sum_{i=1}^{I} \mu_i^2 + 2a_2}{(2 + I)} \tag{A-3}$$

where I is the total number of random effects to be drawn.

A-2.2 Demand for Cereals

Draw of b

b is drawn from a Normal with mean $\bar{\beta}$ and variance/covariance matrix $\frac{W}{N}$), where $\bar{\beta} = \frac{1}{N}\sum \beta_i$.

Draw of W

In iteration t, W is drawn from $IW(2+N, \frac{2*I+N*S}{2+N})$, where $S = \frac{1}{N}\sum(\beta_t^i - b_t)(\beta_t^i - b_t)'$.

Draw of β^i

Each β^i's should be drawn from its conditional distribution

$$\pi(\beta^i|y_i, \xi_{it}) \propto L(y_i|\beta, \xi_{it})\phi(\beta^i|b, W) \tag{A-4}$$

Drawing directly from this posterior is not complicated; therefore draws are approximated through a M-H algorithm. In iteration t, a candidate value $\tilde{\beta}_t^i$ is drawn from a normal distribution. Then I compute the ratio

$$\mathbb{K} = \frac{L(y_i|\tilde{\beta}_t^i, \beta'\xi_{it})\phi(\tilde{\beta}_t^i|b, W)}{L(y_i|\beta_{(t-1)}^i, \beta', \xi_{it})\phi(\beta_{(t-1)}^i|b, W)} \tag{A-5}$$

If $\mathbb{K}$ is greater than a number randomly drawn from a standard Uniform distribution, then $\beta_t^i = \tilde{\beta}_t^i$. Otherwise, $\beta_t^i = \beta_{(t-1)}^i$.

Draw of β'

The vector β' should be drawn from the conditional density

$$\pi(\beta'|y_{it}, \xi_{it}) \propto L(y|\beta, \xi_{it}) \tag{A-6}$$

Once again, sampling from this distribution is not easy. Therefore, I adopt a Metropolis-Hastings procedure. The vector of coefficient is drawn from a normal distribution and the draws are then accepted or rejected in a fashion similar to the one explained for the draws of the β^i's above.

Draw of ξ_{it}'s

Draws of the time-individual specific shocks to the valuation of unknown brands are made jointly with the simulation of the residuals from the channel selection probit. Standard Choleski transformation allows to draw from their joint distribution, specified in (6). These draws of ξ_{it}, however, have to be considered as trial values whose acceptability has to be assessed through updating with the data. In particular

$$\pi(\xi_{it}|y_{it}, \beta, \delta, \Sigma) \propto \phi(\xi_{it}|\rho, \sigma_\xi, \nu_{it}) * L(y_{it}|\beta, \xi_{it}) \tag{A-7}$$

where $\phi(\xi_{it}|\rho, \sigma_\xi, \theta_{it})$ has mean $\delta + \frac{\rho}{\sigma_\xi}\theta_{it}$, and variance $1-\rho$. The density in equation (A-7) is used to evaluate draws from the bivariate truncated in (6) in a Metropolis-Hastings routine.

Computation of δ

In every iteration, $\overline{\delta}$ is computed as

$$\overline{\delta} = 1/R \sum_i \sum_t (\xi_{it} - \rho\theta_{it}) \tag{A-8}$$

with R being the total number of trips in the sample.

In the specifications where δ is parameterized as in equation (7), the parameter are recovered through Bayesian regression.

The model is as follows

$$\delta = \alpha_0 + \alpha_1 Internet + \alpha_2 Weekend + o \rightarrow \delta = \alpha' Z + o, \quad o \sim N(0, \sigma_o^2) \tag{A-9}$$

We know that $\delta_{it} = (\xi_{it} - \rho\theta_{it})$, put a normal prior on α and proceed as follows

- Draw σ_o^2 from an inverse chi-square with parameters $R - 3$ (with three being the dimension of the vector of coefficients and R the number of observations) and $s^2 = \frac{(\delta - Z\tilde{\alpha})'(\delta - Z\tilde{\alpha})}{(R-3)}$
- Draw α from $N(\tilde{\alpha}, \sigma_o^2(Z'Z + A)$, where $\tilde{\alpha} = (Z'Z + A)^{-1}(Z'Z\hat{\alpha} + A\bar{\alpha})$

$\hat{\alpha}$ is the vector obtained by plain OLS in the model in A-9, $\bar{\alpha}$ is the initial prior and the matrix A determines the weight put on the prior. By picking $A = 0.1 * I$ I imply that the prior is diffuse.

Draw of Σ

The variance-covariance matrix of the bivariate truncated is drawn from an inverse Wishart distribution with 2+R degrees of freedom and scale matrix

$$\frac{2 * I + R\hat{\Sigma}}{2 + R} \tag{A-10}$$

where $\hat{\Sigma}$ is the variance matrix implied by the draws of θ_{it} and ξ_{it} in that iteration.