Demand expansion and cannibalization effects from retail store entry: A structural analysis of multi-channel demand

Scott Shriver and Bryan Bollinger*

January 7, 2020

Abstract

For a multi-channel retailer, opening a physical store can lead to cannibalization of online sales as consumers' transportation costs decrease. Holding purchase frequency fixed, this channel switching behavior represents an economic loss to the firm if customers buy fewer or lower margin items in retail stores. However, greater proximity to a retail outlet may increase consumer awareness and consideration of the brand, leading to higher purchase frequencies across channels. In this paper, we devise and estimate a structural model to separately identify the effects of customer-to-store distance on channel switching and brand consideration. Our direct utility framework models outcomes at the level of individual purchase occasions and captures heterogeneous preferences for multiple channels and product categories. Our identification strategy exploits within-customer variation in distance resulting from store entry during a period of rapid retail expansion. We find that retail entry induces channel switching behavior but also leads to greater consideration of the brand. Our estimates imply a 10% reduction in retail store distance increases retail channel expenditures by 2.1% and decreases online channel expenditures by 0.6%, resulting in a 0.8% increase in total expenditures. Through counterfactual experiments, we demonstrate that retail expansion can ultimately limit the ability of the firm to price discriminate across channels, since reducing transportation costs weakens the firm's ability to enforce channel-based segmentation schemes. Retail expansion remains profitable in our setting, where revenue gains from the demand-expanding effect of entry on brand consideration exceed revenue losses resulting from consumers switching to the lower cost retail channel.

*Shriver: Leeds School of Business, University of Colorado, scott.shriver@colorado.edu. Bollinger: Stern School of Business, New York University, bbolling@stern.nyu.edu. We thank the Wharton Customer Analytics Initiative (WCAI) for access to the data and seminar participants at the Marketing Science conference, Duke University, the University of California at San Diego, the University of Southern California, the University of Rochester, the University of Toronto, Stanford University, Cornell University, the University of Colorado and the University of Washington for their comments. We are also grateful to Carl Mela, Laura Kornish, Masakazu Ishihara, Avi Goldfarb and Kitty Wang for useful discussions regarding the paper. All errors are our own.

1 Introduction

Major brands increasingly sell directly to customers through a combination of digital (web, mobile, etc.) and physical (retail) channels. While offering multiple channels entails higher operational costs, the potential benefits to firms are many. First, offering multiple channels can increase demand by increasing the likelihood that a channel will match the consumer's needs for any given shopping occasion. For example, the online channel provides consumers the ability to easily search product assortments, whereas the retail channel allows for direct product interaction and consultation from sales agents. Consumers assess these product information differences, as well as differences in transaction costs, product availability and product prices, when determining their preferred channel for a particular shopping occasion. A second benefit of operating multiple channels is that a presence in physical space (or in digital media) can increase consumer consideration (awareness) of the brand, effectively serving as a form of advertising that can affect demand across multiple channels (Wang and Goldfarb, 2017). The importance of the retail channel in this capacity has been recently reaffirmed, as once decidedly pure-play e-tailers such as Amazon.com have opened retail stores, citing their ability to attract a broader set of consumers and raise brand awareness.[1]

A potential drawback of expanding the retail footprint is that it facilitates channel substitution. Customer switching from online to in-store purchases is particularly concerning for firms that discount merchandise in the retail channel, a common practice among multi-channel firms that sell directly to consumers.[2] To the extent that customers buy lower margin or fewer total items in retail stores, the firm will suffer an economic loss. This empirical setting mirrors the theory model of Yoo and Lee (2011), who show that a monopolist operating online and retail channels charges higher prices in the online channel. More generally, retail expansion can undermine the firm's ability to price discriminate across channels (Venkatesh and Chatterjee, 2006; Zettelmeyer, 2000). differentiation Intuitively, as channels become more equally accessible (less cost separated), channel-based segmentation schemes become less enforceable. Thus, to assess whether a retail expansion is a profitable strategy for existing customers, it is necessary to isolate the effects of entry on both the increase in brand consideration (which drives purchase frequency), as well as the effect on channel expenditures (through channel switching, product purchases, etc.) that result from reducing the costs of

[1] http://www.wsj.com/articles/amazon-rips-page-from-rivals-offline-playbook-1454708011?mod=djem_jiewr_ES_domainid

[2] Unlike in settings with competition between firms who operate in different channels, prices of experience goods in physical stores often exceed those found online when the product is sold in both channels by the same firm. Consumers receive value from transacting in the offline channel, and thus those consumers willing to transact online are those with strong preferences for the brand. This selection effect should lead to lower price sensitivity for consumers transacting online, which is indeed what we find, implying lower optimal markups in the offline channel.

transacting in the retail channel.

To investigate the benefits and risks of operating multiple channels, we develop a unified direct utility framework to measure and decompose the multiple effects of retail store entry. Our model extends the frameworks of Bhat (2008) and Kim et al. (2002) to explain a comprehensive set of demand outcomes, including the frequency with which consumers shop, whether they buy from the online or retail channel, and how much they purchase in multiple product categories. We allow retail entry to affect purchase frequency by relating the arrival rate of brand "consideration events" (potential purchase occasions) to the distance between the consumer's home and the firm's nearest retail outlet. Upon arrival of a consideration event (which implies a need for products), the consumer chooses which channel to visit (if any) given her *expectations* of how she will allocate her shopping budget across the product categories upon visiting a channel. Since traveling to a physical store is costly, she will only do so when the relative benefits of shopping in the physical store from expected product purchases outweigh the costs (Huang and Bronnenberg, 2018; Bronnenberg, 2018). If she chooses to visit a channel, product category shocks are realized and the shopping budget is allocated in accordance with realizations of the category unobservables.

We estimate the model using purchase histories of approximately 10,000 randomly selected customers from a firm that sells semi-durable experience goods exclusively through its own online and retail channels. Customer home locations are observed at the Census block level, yielding highly accurate measures of the distance between where each consumer lives and the nearest retail store. Critically, we observe significant individual-level variation in retail store distance because the firm doubled its retail footprint over our two-year observation window. We leverage within-customer variation in store distance and demand outcomes to identify the effects of retail proximity on purchase frequency and channel expenditure patterns, which effectively controls for time-stationary factors endogenous to store entry. Our model also includes rich time period fixed effects and market-level trend variables to control for time-varying unobservables correlated with entry.

Consistent with the market-level analysis of Wang and Goldfarb (2017), we find that closer retail proximity can lead to increased demand in both channels. Our individual transaction-level model identifies the source of this positive spillover effect from retail entry as consumers' enhanced brand awareness and increased consideration rates of the brand for product needs. Our model also allows for the product category utility in one channel to be affected by recent purchases in the other, which is a second potential source of channel spillovers. However, we find limited evidence of this second type of channel spillover effect,

with only recent retail purchases in a product category diminishing utility for current retail purchases in the same category. Further, we find no substantial spillover of transportation cost savings (from closer retail proximity) into higher retail product expenditures. We are able to assess if transportation cost savings are fungible with consumers' product shopping budgets or separate mental accounts (Thaler, 1985) because our model allows store distance to enter both the shopping budget constraint *and* retail channel utility.

We quantify the economic impact of changing retail proximity (distance) on demand through a series of empirical measures. Channel revenue elasticities with respect to store distance provide a convenient impact metric on aggregate demand for the existing customer base.[3] At the empirical distribution of customer/store locations, we find a total (retail plus online) expenditure elasticity of 0.08 – which implies a 10% reduction in retail store distance increases existing-customer expenditures by 0.8%. The total expenditure increase is driven by an increase in the total purchase frequency (0.9%), which is attributed to greater brand consideration from the closer retail presence. Comparable expenditure elasticity measures by channel imply that decreasing store distance by 10% increases retail channel expenditures purchase by 2.1%, while online expenditures decline by 0.6%. At average expenditure levels, the 10% store distance reduction translates to a 76 cent increase in retail revenues per customer per quarter and a 20 cent reduction in comparable online revenues, a net increase of 51 cents. These results indicate that, at the observed distribution of store distances, revenue increases due to increased consideration are larger than declines in revenues from consumers switching to the retail channel, which has lower average prices.

Through counterfactual simulations, we show that retail expansion can limit the ability of the firm to price discriminate across channels. Specifically, we find that under a uniform percentage price change to all channel products, cross-channel price elasticities increase as retail store distances decline, which indicates consumers' increased willingness to substitute across channels. We also show that with declining store distances, optimal markups increase in the retail channel (with initially lower prices) and decrease in the online channel (with initially higher prices), leading to less price discrimination. However, the counterfactual experiments confirm that demand expansion effects from increased brand consideration dominate losses due to the inability of firms to price discriminate, again highlighting the critical importance of a retail presence as a form of advertising.

The rest of the paper is organized as follows. In Section 2, we discuss the related empirical literature.

[3] In this study we focus on changes in demand patterns among existing customers. Our revenue elasticity measures therefore exclude effects of retail entry on new customer acquisition.

In Section 3 we describe the data used for the study and explore variation in key relationships. In Section 4 we develop our structural model. Section 5 describes our estimation method while model estimates are presented in Section 6. In Section 7, we demonstrate how our results may inform channel pricing policies and retail expansion strategies through counterfactual experiments. We summarize our findings and propose further avenues of research in Section 8.

2 Related literature

Multi-channel retailing has become increasingly popular in recent years as once strictly online retailers such as Amazon have opened traditional (physical) retail outlets, while others such as Bonobos (clothing) and Endy (mattresses) have opened showrooms, which facilitate information gathering and product trial but do not fulfill orders on-site. Neslin et al. (2006) survey the broader literature on multi-channel customer management and discuss how different channel formats can influence consumer behavior. More recently, Lemon and Verhoef (2016) explore the role that channels play in the customer's brand experience and the managerial challenges to creating a coordinated channel strategy. These papers and related works (e.g., Bell et al., 2017; Soysal and Zentner, 2014; Dzyabura et al., 2019) emphasize the retail channel's unique role in enabling consumers to (immediately) assess product fit.

To evaluate the benefits and risks of a multi-channel strategy, it is necessary to conceptually understand and empirically identify the multiple effects that can result from the introduction of a new channel. Much of the literature on new channel introduction dates to the rapid expansion of e-commerce in the late 1990s and early 2000s, and therefore focuses on the impact of adding an online (website) channel to existing retail or catalog channels (e.g., Ansari et al., 2008; Biyalogorsky and Naik, 2003; Deleersnyder et al., 2002; Geyskens et al., 2002; Van Nierop et al., 2011). Researchers have studied the effect of channel introduction in multiple industries including newspapers (Deleersnyder et al., 2002; Gentzkow, 2007), music (Biyalogorsky and Naik, 2003) and apparel (Ansari et al., 2008). Generally, these papers find that the addition of a website does not cannibalize existing retail or catalog channels, due to demand expansion and positive spillovers across channels.

With most firms now operating websites, channel strategy is increasingly focused on when and where to introduce retail outlets. Given the asymmetric nature of online and retail channel experiences, the effects of adding retail to existing online channels could differ from previous studies that explored the opposite

sequence (adding online to existing retail). A number of recent empirical papers, including ours, have explored this issue. For example, Pauwels and Neslin (2015) use vector autoregression (VAR) to simultaneously model multiple demand outcomes, including the frequency and size (in $) of orders, returns and exchanges by channel (online, catalog, store) as well as the total number of customers in the market. The authors find physical store introduction: i) cannibalizes catalog sales but not online sales, and ii) increases the rate of new customer acquisition. However, the VAR approach is silent on the behavioral mechanisms that drive the results and cannot be used for prediction under materially different counterfactual conditions.

Other studies have investigated the effects of adding retail to existing online channels using difference-in-difference methods, typically using geographically aggregated (market-level) data. Since retail entry will only affect consumers within a certain geographic area, difference-in-differences (diff-in-diff) methods can be used to separately identify the effect of new channel introduction from other macro trends, by comparing areas with and without retail store entry. Using market-level panel data from an apparel/home furnishings retailer and a propensity scoring algorithm to select appropriate control markets, Avery et al. (2012) find that catalog and online sales are cannibalized in the short run by retail entry, but that these cannibalization effects diminish over time. Using similar methods, Wang and Goldfarb (2017) investigate the net impact of retail entry on online and existing retail sales in markets defined by Census tracts. They find that retail entry drives sales in both the online and retail channels, particularly in areas with limited ex-ante online sales. Bell et al. (2017) study the effect of offline showrooms and find that showroom entry increases online demand. Similar to Wang and Goldfarb (2017), the authors interpret this spillover as one that results from increased brand consideration.[4]

Two issues complicate previous studies of retail introduction using diff-in-diff methods and market-level data. The first is that it is necessary to identify a suitable control group. Retail entry locations are endogenous to changing demand conditions in that area, and thus it is necessary to find areas with similar demand conditions where the retailer chose not to enter. For example, Bell et al. (2017) define the area of influence of a showroom as a 30-mile radius from the location of the showroom. ZIP codes within the 30-mile radius are in the treated group and those outside of it are in the control group. If consumers in the control and treatment groups are not comparable, the required assumption of parallel trends across groups (in the pre-treatment period) may be violated. The second issue is that market-level treatment effects recovered

[4]Forman et al. (2009) also study the impact of retail stores on online sales, but in the context of a competitive environment. They use a differences-in-differences approach and find a significant decrease in sales on Amazon.com when local bookstores open.

from diff-in-diff estimation measure the net effect of entry on all aspects of the customer's path to purchase, under the observed market configuration. Separate identification of channel entry effects on different stages of the consumer's path to purchase has been highlighted in recent papers as an important area of research (Neslin and Shankar, 2009; Verhoef et al., 2015). Moreover, while market-level treatment effects are of interest, we would not necessarily expect the net entry effect to remain the same in different counterfactual environments. For example, the net effect of retail entry likely depends on a firm's online shipping policy – diff-in-diff effects measured when shipping fees are present are unlikely to apply to contexts where shipping is free.

Structural utility-based models estimated with rich consumer-level data aid identification of channel-related economic primitives and forecasting under unobserved policy regimes. For example, Li and Kannan (2014) use a utility framework to study different stages of the consumer purchase process in an online setting, in which purchases are all made through the website but communication can occur through search and email as well. We similarly let the expected utility of product purchases affect the channel visit utility, although we use a discrete-continuous model in the spirit of Bhat (2008) to incorporate quantity variation in multi-category purchases. We leverage within-customer variation in retail store distance that results from store entry to identify effects of interest. Although store entry locations are endogenous to the types of consumers nearby and to market-level trends affecting those consumers, we can control for these factors using customer-specific fixed effects and market-level time trends. By combining this quasi-experimental variation with a structural model, we can decompose the multiple effects of retail entry. The structural model then allows us to conduct policy simulations under market configurations not observed in the data.[5]

Our structural estimates also allow us to assess the degree to which retail entry undermines the ability of the firm to price discriminate across channels, as suggested in the theoretical literature. Using simulations, Venkatesh and Chatterjee (2006) show that heterogeneity in channel preferences can lead to successful price discrimination strategies across physical and online variants of magazines. Yoo and Lee (2011) explore conditions under which prices will increase or decrease under different channel management structures.

[5]The setup of our model also links it to the literature on store choice, in which consumers similarly consider both the shopping format and the purchase of baskets of goods. Bell and Lattin (1998) study the choice of retail store format (EDLP or HILO) in an environment with different levels of price uncertainty. The choice of channels in our context has a similar character, in that channel formats differ in terms of their ability to provide product information. The store choice literature has also emphasized the importance of planned expenditure levels (or "basket size") and shopping trip fixed costs (e.g. Bell, Ho, and Tang, 1998) as determinants of shopping format choices. We similarly model channel choice as a function of the shopping trip budget and channel transaction costs. Whereas the standard practice in the store choice literature has been to treat expenditure levels as exogenously determined (Bell and Lattin, 1998; Bell et al., 1998), we endogenize the consumer's expenditure decision (through category quantity choices) as well as her choice of channel.

They show that with vertically integrated channels, prices for the retail channel will increase (towards the online channel) with increased substitution across retail outlets. Finally, Kireyev et al. (2017) demonstrate that cross-channel price-matching (where firms guarantee to match the lowest price across their own channels) can dampen competition between the online and retail channel, and thus facilitate price discrimination strategies.

From a methodological standpoint, we expand upon the discrete-continuous frameworks of Bhat (2008), Kim et al. (2010) and Thomassen et al. (2017). Our framework is unique in that it jointly models path-to-purchase outcomes as a multi-stage decision process with separate unobservables realized at each decision stage. Our second-stage category quantity model follows Bhat (2008), which we extend by incorporating a first stage channel choice decision that is in part based on the expected utility consumers will get from optimally allocating their shopping budgets upon visiting a channel. This multi-stage formulation requires us to develop a computationally efficient method to integrate out the shocks (and associated optimal quantities) that are unobserved to consumers before visiting the channel. Kim et al. (2010) jointly model cell phone usage and plan choice as a function of expected usage in a one-stage model with a shock structure that does not account for differences in expected and actual usage. Thomassen et al. (2017) allow store and quantity choices to be jointly determined, but the product utility shock is additive in nature and thus does not directly affect quantities purchased.[6]

3 Data

3.1 Sources

The primary data for the study come from a North American speciality retailer that sells to customers exclusively through its e-commerce website (the online channel) and network of retail stores (the retail channel).[7] The brand sells a variety of apparel products including clothing, footwear and accessories. While fewer than 0.1% of products are offered exclusively via one channel, operational issues such as stock-outs may result in unobserved assortment variation across the channels.[8]

[6]Our paper also differs from Thomassen et al. (2017) in that we model outcomes at the level of individual purchase occasions rather than weekly-level aggregates, allowing us to recover structural estimates of consumer transaction costs. Further, our MCMC estimation procedure allows for a richer heterogeneity specification than would be feasible with simulated GMM.

[7]Confidentiality agreements preclude us from disclosing the identity of the brand or revealing precise descriptions of the products in their portfolio. Note also that exclusivity implies the firm's products are not available through other retailers.

[8]Lacking data on store and online channel inventories, we cannot characterize such assortment variation precisely. In general, we expect fewer stock-outs in the online channel, which operates from large logistics facilities.

In raw form, the firm data are tables exported from a relational database. The relevant tables and fields for our analysis are as follows:

1. Customer metadata (random sample) – home location (Census block), age, first purchase date.
2. Customer purchase histories – date, channel, SKU quantities/prices/return indicators for each transaction.
3. SKU metadata – SKU to product category mapping.
4. Store metadata – latitude/longitude, open/close dates.

The first table contains a randomly generated list of 14,000 customers with home location and demographic information, while the second table contains the complete purchase histories of those customers from July 2010 through June 2012. Purchase transaction records indicate the date, channel format, the quantities/prices of individual SKUs purchased, and a return indicator for each SKU. SKU metadata in the third table include the product category assignment (as well as fields such as size and color), which we leverage to aggregate demand outcomes to the category level.

The final table from the firm contains retail store metadata, including exact entry dates and locations. From this information and knowledge of a customer's Census block, we can calculate the distance between her home and the brand's nearest retail store at any point in time with high (~1/3 mile) accuracy. Critically, we observe significant within-subject variation in this retail distance measure due to the entry of new stores – during our two-year period of study, the brand expanded its retail footprint from 37 retail outlets to 75. We explore the effect of retail distance on various demand outcomes in Section 3.4, with emphasis on using within-subject variation to identify the relevant effects.

We augment the firm data with two external data sources:

1. Census block group demographics (US Census) – We collect market characteristics (median income, commute times, rural population proportion, etc.) for the complete set of Census block groups in the US. We impute market characteristics to customers using the parent block group of the customer's home Census block.
2. Tax rates by zip code and channel (avalara.com) – We collect sales tax rates by US zip code and use information from the firm's website (the list of states for which the firm collects online sales tax) to impute tax rates to customers by channel. The customer's retail channel tax rate is the rate in the zip code containing the brand's nearest retail outlet. If the customer resides in a state with online tax

collection, her online channel tax rate is the tax rate in the zip code containing her home (and is equal to zero otherwise).[9]

3.2 Preparation

To facilitate our empirical analysis, we refine and structure the estimation sample along three important dimensions. First, as a conservative way to remove consumers who have no meaningful tradeoff between using the online and retail channels, we restrict our sample to US customers who live within 500 miles of one of the brand's stores at some point during the observation window. The restricted sample contains 10,239 customers and a total of 29,095 purchase transactions.

Second, we aggregate continuously-varying price and store distance information with summary discrete-time measures at the quarterly level. By contrast, variables capturing state dependence (described in Section 4.2) are continuously updated in response to observed purchase transactions.

Third, we abstract from SKU-level choices and characterize purchase outcomes in terms of: a) purchase channel indicators, and b) purchase quantity indices associated with the six top-level product categories in the brand's SKU database. We focus on product categories because the perpetual introduction and retirement of seasonal apparel products makes individual SKUs both conceptually and computationally impractical as the unit of analysis, given our objective to predict demand both in and out of sample. These category definitions are sufficiently broad so as not to be season-specific (e.g., a footwear category would encompass winter boots as well as summer sandals) or subject to product-specific fashion cycles (Yoganarasimhan, 2012, 2017)while still narrow enough to share common requirements regarding fit assessment (e.g., the need to try on footwear does not vary greatly across shoe styles but is presumably different from the need to try on pants). As our principal interest lies in the net economic contribution of channels on demand, we omit returned items from our computation of category purchase quantities.[10]

3.3 Summary

We first discuss variation in the data at the customer-quarter level, and then turn to discussion of variation at the level of individual purchase occasions. Table 1 summarizes variables with quarter-level variation

[9] Retail sales tax rates downloaded from https://www.avalara.com/products/taxrates/. Mapping between customer locations (Census blocks) and zip codes is accomplished using a crosswalk file generated from the MABLE/Geocorr12 Geographic Correspondence Engine at http://mcdc.missouri.edu/websas/geocorr12.html.

[10] Approximately 15.0% of purchased items are returned, corresponding to 11.4% of total expenditures.

(number of purchases, quarterly expenditures by channel, retail store distance) and with cross-sectional variation (demographics, taxes). Due to acquisition of new customers, the data summarized in Table 1 is unbalanced across the 8 quarters of study and 10,239 individuals. We note that approximately half of these individuals make purchases in both channels, implying than channel switching is not uncommon behavior.

From Table 1, we see that on average customers make 0.5 purchases per quarter (2 purchases per year). As seen in the full distribution of purchases per quarter (Figure 1 below), some customers are very frequent shoppers, with more than 10 purchases per quarter, suggesting a strong need for heterogeneous effects relating to purchase frequency. Average expenditures per customer per quarter (aggregating over purchase occasions) are approximately $30 in the online channel and $34 in the retail channel. Our retail distance measure implies customers live approximately 43 miles from the nearest store. Temporal variation in the retail distance distribution, which will be used to estimate retail distance effects on brand consideration, retail channel utility and retail channel transportation costs is shown in Figure 2. In the figure, we plot the distributions of customer retail distance (in log scale) for the initial and final periods, which shows that retail distance declines appreciably over time. Further, the within-customer variation induced by observed retail entry events is substantial, with a standard deviation of 28.4 miles.

The cross-sectional data include directly observed demographic variables (age), demographic variables imputed from matching Census 2010 block groups, and customer location-specific sales tax rates for online and retail purchases. While the means strongly suggest the firm's core market is affluent, middle-aged women, there is also substantial variance in the demographic variables. As the near equivalence of the online and retail mean tax rates suggests, for the vast majority (97%) of consumers, online and retail tax rates are equal. Differential rates generally apply for those individuals living near state boundaries, where their home and closest retail outlet are located in different tax jurisdictions.

Panel variables		observations	mean	std dev	min	max
	quarter	57,554	5.130	2.208	1	8
	purchases (#/qtr)	57,554	0.506	1.005	0	36
	online expenditures ($/qtr)	57,554	30.218	136.825	0	8181.450
	retail expenditures ($/qtr)	57,554	33.989	120.613	0	5532.000
	retail distance (mi)	57,554	42.656	81.436	0.1	499.741

Cross sectional variables		observations	mean	std dev	min	max
Customer characteristics	age	10,239	40.460	11.754	18	94
Market characteristics	rural	10,239	0.115	0.284	0	1
	white	10,239	0.800	0.186	0	1
	college	10,239	0.548	0.213	0	1
	commute time (min)	10,239	28.140	7.512	8.628	61.746
	median HH income (k)	10,239	97.604	49.646	2.499	250.001
	online tax rate	10,239	0.072	0.017	0	0.096
	retail tax rate	10,239	0.074	0.014	0	0.096

Table 1: Customer/quarter panel summary statistics

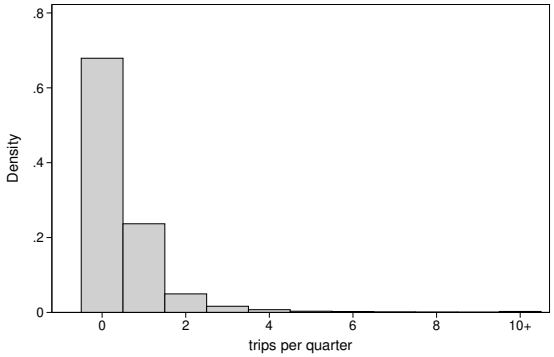


Figure 1: Frequency of purchase occasions per quarter

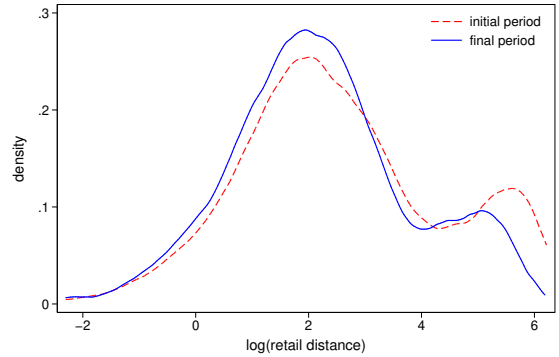


Figure 2: Customer retail (log) distance distributions in initial and final periods

Price indices, which are not specific to individuals but vary by time, category and channel, are summarized graphically in Figure 3. Similar to previous treatments in the marketing and economics literatures (e.g., Gordon et al., 2013; Chevalier et al., 2003), we compute category (k) and channel (c) specific price indices for a quarter (t) as geometric expenditure-weighted means of SKU-level prices:

$$p_{kct} = exp\left(\frac{\sum_{j\in J_t,k} w_j \, log(p_{jct})}{\sum_{j\in J_t,k} w_j}\right) \quad where: \; w_j = \frac{\sum_c \sum_t e_{jct}}{\sum_j \sum_c \sum_t e_{jct}} \tag{1}$$

In (1) above, p_{jct} is the average (across consumers) price paid for SKU j in channel c and quarter t, while e_{jct} is the total expenditure (summed over consumers) for the same SKU. We use J_t to represent the set of SKUs offered in quarter t. The applied weights, which are the fraction of total observed expenditures

attributable to a given SKU, naturally reflect more popular products in the resulting index. As is apparent from Figure 3, for most categories, prices are slightly lower in the retail channel. A SKU-level analysis of prices reveals that these differences stem from more aggressive discounting in the retail channel.

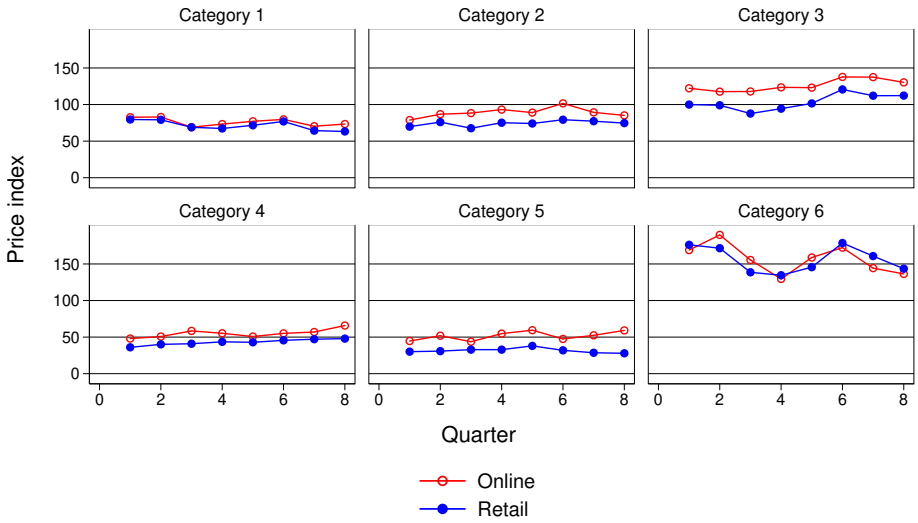


Figure 3: Channel-specific category price indices

Next we turn to data on individual purchase occasions. We summarize observed purchases in terms of the selected channel (where we code online as 1 and retail as 2), the total expenditure, and the expenditure shares for each of the six product categories. The purchase transaction data summarized in Table 2 indicates that the average per-purchase expenditure is approximately $127 and that the 57% of observed purchases are in the retail channel. Expenditure shares tend to be highest for category 1 and lowest for category 6, which is the highest priced category.

	observations	mean	std dev	min	max
channel	29,095	1.57	0.50	1	2
expenditure	29,095	127.01	125.42	1	1245.95
share 1	29,095	0.32	0.41	0	1
share 2	29,095	0.12	0.29	0	1
share 3	29,095	0.13	0.30	0	1
share 4	29,095	0.20	0.35	0	1
share 5	29,095	0.18	0.34	0	1
share 6	29,095	0.05	0.21	0	1

Table 2: Purchase occasion summary statistics

We examine the relationship between expenditures and channel formats in Figure 4, which plots kernel density estimates of trip expenditures by channel. It is clear that customers spend more per trip in the online

channel (average values are \$138.03 for online and \$118.59 for retail), and that there are a large number of small expenditure transactions in the retail channel. We explore variation in category expenditure shares by channel in Figure 5. The bulk of the expenditure share variation across channels is linked to categories 1, 5 and 6, where there is a clear preference for category 6 in the online channel and for categories 1 and 5 in the retail channel.

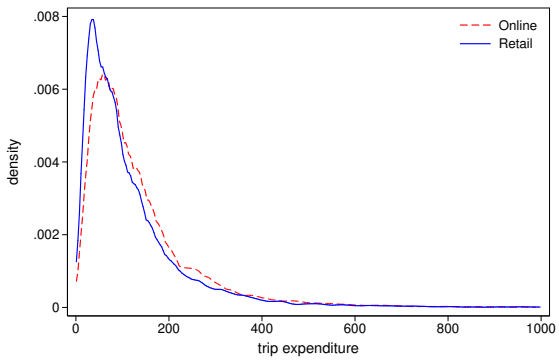


Figure 4: Expenditure distributions by channel

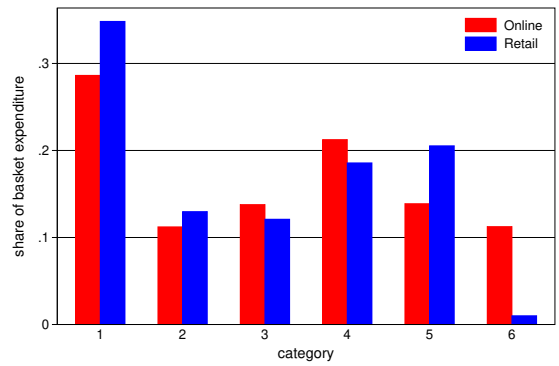


Figure 5: Average expenditure shares by category and channel

3.4 Descriptive analyses

In this section, we leverage the entry events to explore how distance to the retail outlet affects demand along three dimensions: purchase frequency, the channel chosen for purchase, and the total expenditure per purchase.

Throughout this section and in our model development, we define the variable d as the distance to the nearest retail outlet.

3.4.1 Purchase frequency

To further quantify the effect of distance on brand consideration, we model the number of purchase occasions for customer i in quarter t, L_{it}, as a Poisson arrival process and use the customer/quarter panel data to estimate the model. The specification includes bi-level (individual, quarter) fixed effects (δ_i, μ_t) to control for unobservables – the effect of retail store distance (α) is thus identified by deviations from individual-specific average purchase incidence rates, after controlling for time trends common to all individuals. Formally, the specification is:

$$L_{it} \sim Poisson(\rho_{it}),\ log\,(\rho_{jt}) = \alpha\, log(d_{it}) + \delta_i + \mu_t \tag{2}$$

The estimate of α from this regression is $\hat{\alpha} = -0.104$ with a standard error of 0.016, which is significant below the 1% level. Given the log-log specification for ρ and d, we may interpret $\hat{\alpha}$ as an elasticity, so that a 10% decrease in retail store distance corresponds to a 1.0% increase in the mean number of purchases per quarter. We conclude that the increased proximity to a retail outlet (smaller store distance) has a significant positive impact on shopping incidence rates.

We extend our analysis by considering purchase frequency by channel, which we summarize in Figure 6. Plots on the left-hand side of Figure 6 graph the conditional mean number of purchases per quarter (with 95% confidence intervals) for the retail (panel a) and online (panel b) channels at various distances from the nearest retail store (in addition to the smoothed conditional mean estimates). Whereas the left-hand side plots are raw data, the right-hand side plots display predicted values from regressions of equation (2), where the $log(d_{it})$ term is replaced with a semi-parametric specification based on 6 contiguous distance bands. These plots show that not only do retail purchases increase in frequency with shorter distances, so do online purchases, indicating that retail channel proximity increases awareness and/or consideration of the brand. However, compared to retail purchases, increases in online purchases only occur in quite close proximity to retail locations.

3.4.2 Channel choice

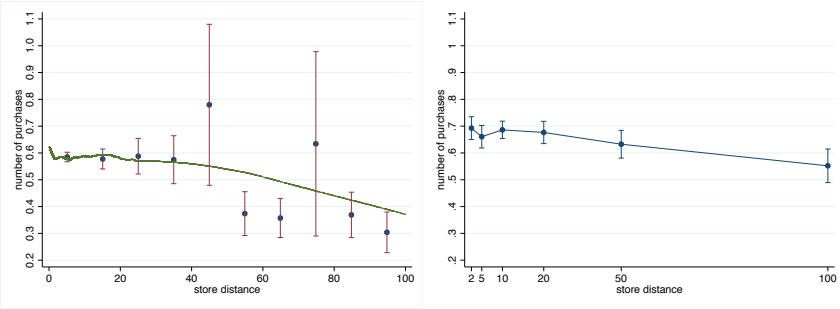
Next we analyze the effect of retail distance on channel format choices. We denote customer i's channel choice for the l'th purchase in quarter t as c_{itl}, and following our previous convention, associate a channel choice of 2 with the retail channel. We proceed by modeling the probability of a retail channel choice using a binary logit model, again including bi-level fixed effects:

$$Pr(c_{itl} = 2) = \frac{exp\,(\alpha\, log(d_{it}) + \delta_i + \mu_t)}{1 + exp\,(\alpha\, log(d_{it}) + \delta_i + \mu_t)} \tag{3}$$

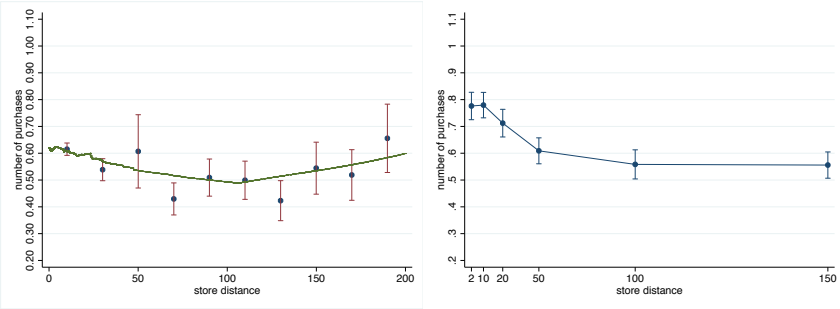
The estimated α is $\hat{\alpha} = -0.531$ with a standard error of 0.064, which is significant below the 1% level. Unsurprisingly, proximity to a retail outlet has a strong effect on channel choices – the closer to the retail outlet, the higher the probability of a retail (vs. online) purchase. The intuitive interpretation of this result is that lower transportation costs or higher retail outlet utility due to store entry increase use of the retail channel. As before, Figure 7 shows the relationship between this dependent variable and distance in the raw data as well as the regression estimates when using a semi-parametric function of distance. Collectively, these analyses provide robust evidence of channel choice dependence on retail distance.

Figure 6: Distance effect on # of customer purchases per quarter

(a) Retail

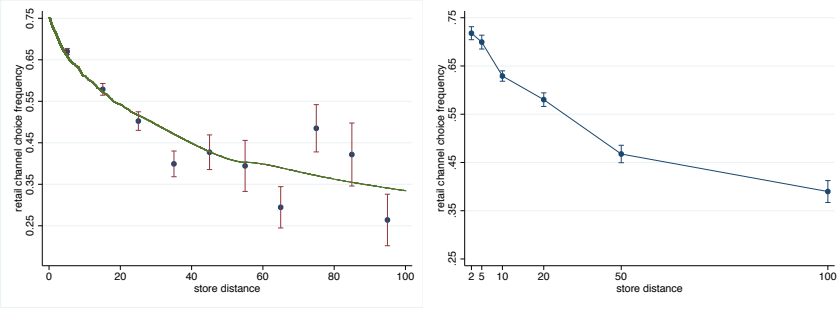


(b) Online



Notes: Left-hand plots are raw data, while right-hand plots are predicted values from semi-parametric regressions of equation 2. Error bars designate 95% confidence intervals. In the right side graphs, upper values of distance bands are displayed on the axis.

Figure 7: Distance effect on retail channel choice frequency



Notes: Left-hand plot is raw data, while the right-hand plot is predicted values from a semi-parametric regression of equation 3. Error bars designate 95% confidence intervals. In the right side graphs, the upper value of the distance band is displayed on the axis.

3.4.3 Expenditure level

Although retail outlet proximity increases customer purchase frequency, to infer the net effect on total expenditure levels with the brand, we must also consider the potential impact of retail distance on the expenditure per purchase. In a unified direct utility model, transportation costs avoided by the consumer after a more proximate retail store enters may be partially used to purchase more of the inside goods, but only if the transportation budget is fungible with the budget used for the categories we study. To explore this issue, we use the purchase occasion data and model the expenditure for the l'th purchase by customer i in quarter t, e_{itl}, using a log-log model with bi-level (individual, quarter) fixed effects (δ_i, μ_t):

$$log(e_{itl}) = \alpha\, log(d_{it}) + \delta_i + \mu_t + \varepsilon_{itl} \tag{4}$$

The estimate of α from this regression is $\hat{\alpha} = -0.003$ with a standard error of 0.014. The sign of $\hat{\alpha}$ is consistent with expenditures increasing at closer retail distances; however, the effect is not significant at any conventional level. Similar results hold if we condition upon only online or only retail purchases.

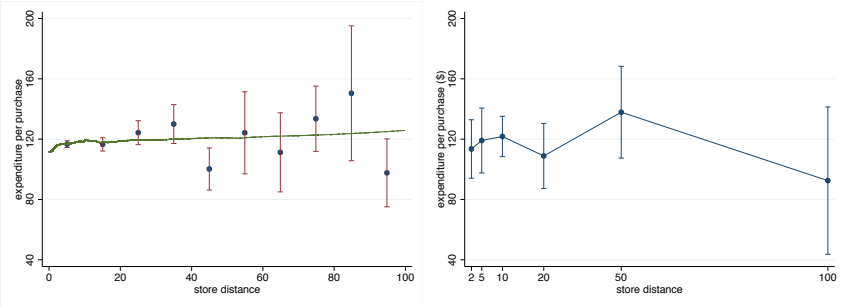
We plot the raw data (average expenditure versus store distance) in the left side graphs of Figure 8 and plot predictions from semi-parametric variants of equation (4) in the right side graphs. These graphs suggest no obvious relationship between expenditure per purchase and distance. Although our structural model will allow expenditure levels to depend on retail distance through transportation costs (which would reduce the budget available for retail purchases), this descriptive evidence suggests limited scope for monetary transportation costs being fungible with shopping budgets. Therefore, our model also needs to be able to capture transportation costs when consumers treat them as coming from a distinct budget (consistent with mental accounting).

4 Model

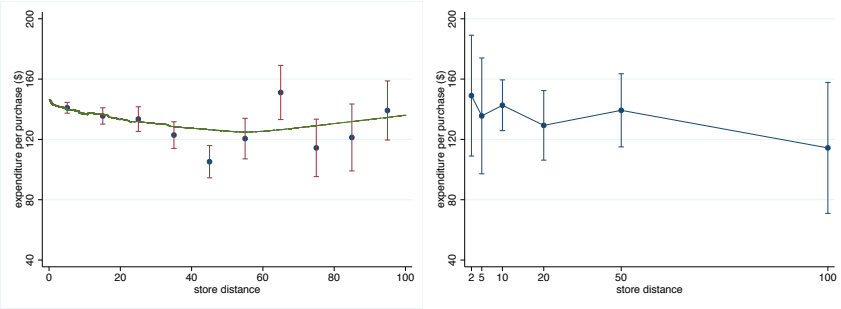
Our modeling objective is to use economic primitives to explain and predict a comprehensive set of demand outcomes: how frequently consumers purchase, which channels they use to purchase, and how much they spend across multiple product categories. In accordance with this objective, we model demand at the level of individual purchase occasions. To facilitate the exposition, we first develop the model for homogeneous consumer preferences. We subsequently incorporate heterogeneous preferences into our empirical specifications, as discussed in Section 4.3.

Figure 8: Distance effect on customer expenditure per purchase

(a) Retail



(b) Online



Notes: Left-hand plot is raw data, while the right-hand plot is predicted values from a semi-parametric regression of equation 4. Error bars designate 95% confidence intervals. In the right side graph, the upper value of the distance band is displayed on the axis.

Figure 9 below outlines our proposed model, which conceptualizes demand as originating with a brand consideration arrival process. Specifically, the number of times a consumer *considers* purchasing from the focal brand in a given period is assumed to follow a Poisson distribution with (quarterly) rate parameter μ. The consideration arrival process may be conceptualized as reflecting: (i) Poisson process arrivals of the consumer's need for product, and (ii) the conditional (Bernoulli) probability of considering the focal brand given product need.[11] Note that the consideration arrival process is different than the observed purchase rate in the data because for each consideration event we allow consumers to completely allocate their shopping budget to the outside good. The purchase rate we observe in the data reflects the brand consideration arrival rate and the probability some expenditure is allocated to the inside goods.

Conditional upon a brand consideration event, consumers engage in a two-stage decision process: in the first stage, the consumer makes the channel choice decision, and in the second stage, the consumer allocates her budget (by choosing category quantities) across the six inside-good categories and the outside good. The

[11] In our setting, product need arrival (N) and brand consideration (C) outcomes are not directly observed and hence not separately identified. If N~Poisson(λ) and C~Bernoulli(ζ), the distribution of consideration events Y≡N|C=1 is Poisson($\lambda\zeta$). Our model effectively recovers the compound rate parameter $\mu = \lambda\zeta$. We envision retail store distance affecting the arrival process through the conditional probability of brand consideration given product need rather than the need arrival rate itself.

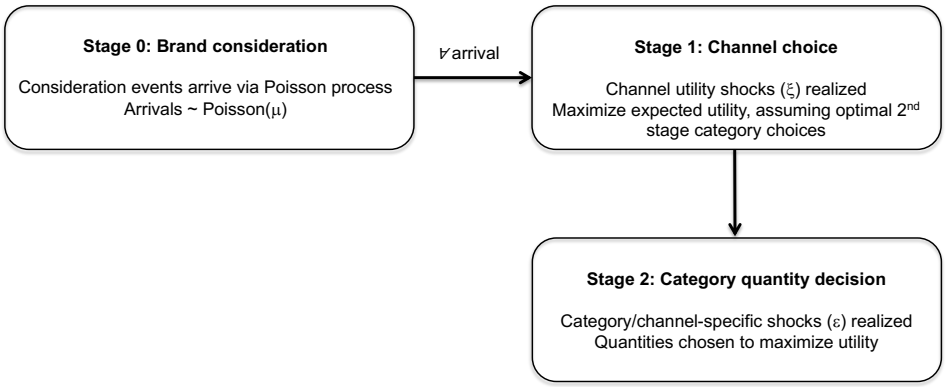


Figure 9: Model overview

consumer is forward-looking in the sense that she anticipates how she will allocate her total shopping budget (b) across the six categories and the outside good upon visiting a channel, if she chooses to do so. However, the channel-specific category shocks are not realized unless she choses to visit the channel. Thus the tradeoff the consumer faces is whether to incur the (fixed) costs of visiting a channel in order to observe the realized shocks (and then optimally allocate the budget to purchase products), or to not visit either channel, in which case the entire budget is allocated to the outside good. The latter condition will obtain if all realizations of the category shocks are sufficiently negative. Category quantity levels are constrained such that the total shopping budget is equal to the sum of: (i) category expenditures (quantity times price, plus applicable sales taxes), (ii) the outside good expenditure, and (iii) channel fixed costs. Channel fixed costs include shipping costs (online) and transportation costs as a function of store distance (retail). Because some or all of the costs associated with visiting the retail channel may be incurred from a separate budget, we also allow the distance to the nearest retailer to enter the utility of selecting the retail channel.[12]

4.1 Consumer choice model

Stage 1: Channel choice decision Conditional upon a brand consideration event, the consumer maximizes her expected total utility, $V(\vec{y})$, over the channel choice vector $\vec{y}$ with length equal to the number of channels, C. The elements of $\vec{y}$ are either 0 or 1, with the latter corresponding to an observed channel choice.

[12]We acknowledge that in some settings consumers may shop in sequences other than the one we describe. For example, consumers may first determine the set of products to be purchased in the form of a shopping list (e.g. Bell, Ho, and Tang, 1998) and subsequently choose the channel from which to shop. In most cases (including ours), the consumer's decision sequence cannot be identified from observable data, requiring any complete model of multi-channel, multi-product shopping behavior to employ comparable assumptions.

The consumer's problem is therefore formulated as:

$$\vec{y}^* = \underset{\vec{y}}{argmax}\ V(\vec{y}) \tag{5a}$$

$$subject\ to: \sum_{c=1}^{C} y_c \in \{0,1\} \tag{5b}$$

$$where\ V(\vec{y}) \equiv \sum_{c=1}^{C} y_c\,(v_c + \Omega_c) + \left(1 - \sum_{c=1}^{C} y_c\right)\log(b) \tag{5c}$$

$$and\ \ v_c \equiv \underset{\Psi_c}{E}\left[\sum_{k=1}^{K} \Psi_{ck}\,\gamma_{ck}\,\log\left(\frac{q^*_{ck}(\Psi_c)}{\gamma_{ck}} + 1\right) + \log(z^*)\right] \tag{5d}$$

$$E\ subject\ to: \left(\sum_{k=1}^{K} q^*_{ck} p_{ck}\right)(1 + r_c) + f_c + z^* = b,\ z^* > 0 \tag{5e}$$

The vector $\vec{y}^*$ is the optimal channel choice that maximizes the expected total utility, the sum of the utility from visiting a channel Ω_c (e.g. utility derived from browsing products in a retail store) and the expected utility from product purchases made at the channel (v_c). At any one occasion, the consumer can only chose one channel, as indicated in equation 5b.

The total utility is given in equation 5c – if neither channel is chosen, the utility is simply equal to the utility from the outside good, $\log(b)$, where b is the trip-level budget.[13] The expected product utility is given by the term v_c in which the expectation is taken over the vector of channel-specific category shocks Ψ_c, whose realizations are observed upon channel visitation. The Ψ terms are the so-called "baseline" marginal utilities (whose empirical specifications are shown in Table 3), because Ψ_{ck} is the marginal utility obtained in the limit of consuming zero quantity ($\underset{q_{ck}\to 0}{lim}\frac{\partial U}{\partial q_{ck}}$). The addition of one to $\frac{q^*_{ck}(\Psi_c)}{\gamma_{ck}}$ inside the $\log(\cdot)$ expression enforces a lower bound of zero for each of the additively separable category/channel sub-utilities. The role of the γ parameters are to alter the rates at which the consumer's utility satiates in purchase quantity – a higher γ implies a lower rate of satiation.

The expectation for product utility is taken subject to the constraint given in equation 5e. Upon visiting a channel, the consumer will allocate the entire budget among the six inside goods (the six categories) and the outside good, after incurring the channel transaction costs (f_c), and accounting for channel-specific sales taxes (with rate r_c). We assume that the product shipping cost is included in f_c for the online channel. For the retail channel, it is not clear whether transportation costs will be treated as a fixed cost that reduces the

[13] Imposing diminishing marginal utility for the outside good z (though the $log(\cdot)$ expression) is essential in our context. Assuming linear utility in the outside good (z) would result in optimal demand conditions that do not involve the shopping budget or channel fixed costs, which are key objects of inference.

budget (b) available for the category, or whether they will be treated as coming from a separate budget, so we include a function of store distance in retail fixed costs, f_c (which treats the budgets as fungible), as well as in the retail channel utility, Ω_c (which treats the budgets as separate). The Ω_c term also captures utility that stems from the shopping activity itself. In practice, we include a utility intercept for the retail channel to allow consumers to prefer one channel over the other, irrespective of the fixed costs and the product utility.

As is standard, some expenditure on the numeraire outside good z is essential ($z > 0$) for each consideration event. z is not directly observed but, as will be seen, its value may be inferred from data and model assumptions. We further discuss computational issues related to calculating expected product utility in Appendix B.

Stage 2: Category purchase decisions in the visited channel If the consumer chooses to visit a channel, she receives the realizations of the Ψ_{ck}. She then optimally allocates her budget across the inside and outside goods. We permit this entire budget to be allocated to the outside good, which can occur if all realizations of the category Ψ_{ck} are small enough. Formally, the consumer's second stage decision may be written as:

$$\vec{q}_c{}^* = \underset{\vec{q}_c, z}{argmax}\, U_c(\vec{q}_c, z) \tag{6a}$$

$$subject\ to:\ z > 0,\ q_{ck} \geq 0,\ z + \left(\sum_{k=1}^{K} q_{ck} p_{ck}\right)(1 + r_c) + f_c = b \tag{6b}$$

$$where:\ U_c(\vec{q}_c, z) \equiv U(\vec{q}, z \,|\, y_c = 1) = \sum_{k=1}^{K} \Psi_{ck}\, \gamma_{ck} \log\left(\frac{q_{ck}}{\gamma_{ck}} + 1\right) + \log(z) \tag{6c}$$

4.2 Empirical specifications

In order to solve the model and derive the corresponding likelihood function, it is necessary to impose functional form and stochastic assumptions on $\{\Psi_{ck}, \Omega_c, f_c, b, \mu\}$. We summarize our empirical specifications for these components in Table 3 below.

	Functional form	Distributional assumptions
Category utilities	$\Psi_{ck} = exp\left(\psi_{ck} + \tilde{\psi} S_k + \varepsilon_{ck}\right)$	$\varepsilon_{ck} \sim N\left(0, \Sigma_c\right)$
Channel utilities	$\Omega_1 = \xi_1, \quad \Omega_2 = \omega_2 + \tilde{\alpha} d^{\tilde{\kappa}} + \xi_2$	$\xi_c \sim EV\left(0,1\right)$
Channel fixed costs*	$f_1 = 7.95, \quad f_2 = \alpha d^{\kappa}$	
Trip budget	$b = exp\left(\tau + \beta X\right)$	
Consideration arrival*	$\mu = exp\left(\delta + \tilde{\beta}\tilde{X}\right)(1+d)^{\zeta}$	

* d = retail store distance

Table 3: Empirical specifications for $\{\Psi_{ck}, \Omega_c, f_c, b, \mu\}$

Category marginal utilities Ψ_{ck} are constrained to be positive through the $exp(\cdot)$ operator, and are comprised of a category intercept (ψ_{ck}), state dependence term ($\tilde{\psi} S_k$), and a stochastic term (ε_{ck}). S_k is a category-specific vector of length four that corresponds to interactions of indicators for the current channel and indicators for prior purchase channels in the same category (online X online, online X retail, retail X online, retail X retail).[14] We assume a structural interpretation for these state dependence terms, namely that the $\tilde{\psi}$ captures the impact of the last purchase occasion on category utility. The ε_{ck} are normally distributed channel/category specific shocks that capture product fit and assortment information. We allow for full covariance between the category shocks within each channel. We expect that shock variances are influenced by: i) unobserved match values of the customer with the channel/category, and ii) unobserved channel-specific product assortment variation (e.g. unobserved stock-outs). Given the exponential form in Table 3, a higher shock variance yields a higher expected utility, such that larger shock variances are associated with higher average channel match values.

The channel utility Ω_c is comprised of an extreme value shock (ξ_c), plus an intercept (ω_c) and utility shifters where appropriate. For the online channel, we normalize the intercept to zero ($\omega_1 = 0$) and do not include any utility shifters. For the retail channel, we include an intercept that captures retail format preferences and a flexible function of the consumer's distance (d) to the nearest retail store, $\tilde{\alpha} d^{\tilde{\kappa}}$, in which $\tilde{\kappa}$ is constrained to be positive. For channel fixed costs, we set $f_1 = 7.95$ because the firm charges a flat rate shipping fee of \$7.95 per online order. Although customer-store distances are known, consumer retail channel fixed (transportation) costs are not fully observed as mileage costs and commuting patterns are unobserved. We therefore capture retail fixed costs using a similar functional form of the consumer's distance to the nearest retail store: $f_2 = \alpha d^{\kappa}$, in which α and κ are both constrained to be positive.

In addition to an intercept (τ), the trip budget specification incorporates additional controls through the

[14] Prior purchase indicators are set to zero if no previous purchases are observed in the current or previous time period (quarter). We therefore assume purchases more than 2 quarters old do not influence current category utility.

shifter matrix X. Our specification of X includes quarter fixed effects and Census combined statistical area (CSA) time trends to control for common unobserved market factors that potentially affect both demand (expenditure patterns) and firm retail entry (and thus, retail store distance).

Finally, the consideration arrival rate parameter μ is specified as a flexible function of retail store distance, an intercept (δ) and additional controls ($\tilde{X}$). The term incorporating retail store distance, $exp\left(\delta+\tilde{\beta}\tilde{X}\right)(1+d)^{\zeta}$, captures the effect of retail store proximity on a consumer's propensity to consider the brand, and is constructed such that the arrival rate is bounded above at $exp\left(\delta+\tilde{\beta}\tilde{X}\right)$ as $d \to 0$ and at zero as $d \to \infty$ (provided $\zeta < 0$, as is expected). In $\tilde{X}$, we again include quarter fixed effects and CSA time trends.

4.3 Heterogeneity specification

We allow for heterogeneous product and channel preferences by incorporating individual-specific heterogeneity in the category (ψ_k) and retail channel (ω_2) utility intercepts in a hierarchical model. We further allow for individual-specific shopping budgets in τ, and brand consideration rates in δ. We collect the heterogeneous parameters in the vector $\theta_i^1 \equiv \{\psi_{ick}, \omega_{ic}, \tau_i, \delta_i\}$. Similarly, we collect the homogeneous parameters in the vector $\theta^2 \equiv \left\{\tilde{\psi}, \gamma, \Sigma_1, \Sigma_2, \alpha, \kappa, \tilde{\alpha}, \tilde{\kappa}, \beta, \tilde{\beta}\right\}$. We assume a hierarchical model such that:

$$\begin{aligned}
\theta_i^1 &\sim N\left(Z_i\Upsilon, \Xi\right) \\
\theta^2 &\sim N\left(\overline{\theta^2}, A_{\theta^2}^{-1}\right) \\
vec(\Upsilon) &\sim N\left(vec\left(\overline{\Upsilon}\right), \Xi \otimes A_{\Upsilon}^{-1}\right) \\
\Xi &\sim IW\left(\upsilon, W\right)
\end{aligned}$$

Heterogeneous parameters θ_i^1 are thus projected onto demographic variables Z_i using the hyperparameters Υ. In Z_i, we include the customer and market demographic variables from Table 1. For the conjugate prior distributions, we assume that the A matrices are $0.01I$, where I is the identity matrix. We also set $\upsilon = length(\theta_i^1)$ and $W = \upsilon I$.

5 Estimation

We use Markov Chain Monte Carlo (MCMC) methods to estimate the model. MCMC methods require evaluation of the model likelihood, which we derive in Appendix A. The likelihood function follows from the consideration (Poisson) arrival process in stage 0, and the consumer optimization problems in stages 1

(channel choice) and 2 (category quantity choices).

Since channel fixed costs are sunk during the second decision stage, the (conditional) utility function and budget constraint equation are continuously differentiable in category quantities. Thus, we can apply the Kuhn-Tucker (KT) theorem to derive optimality conditions for product category-related parameters. For chosen categories, the KT conditions provide a one-to-one mapping between optimal quantities and category shock values. For unchosen categories, the KT conditions provide (upper) bounds on the category shock values. The second stage conditional likelihood therefore comprises a Jacobian term for chosen categories and a region of integration over shock values for unchosen categories.

By contrast, the first stage optimality conditions involve only inequality conditions – namely, that an observed channel choice must correspond to greater utility than from listing alternative channels or from allocating the entire shopping budget to the outside good. When considering channel choices, consumers formulate expectations about optimal category quantities during the second decision stage – i.e., in the first stage category shocks are known only in distribution. Under the assumption of extreme value channel utility shocks (and a non-stochastic outside good utility), the first stage likelihood takes the form of a modified logit expression. Finally, the timing of observed purchases is linked to the consideration arrival process by incorporating the probability of no-purchase given a consideration event into the Poisson arrival likelihood.

Given our hierarchical model specification, we can use the following Gibbs sampler to simulate from the posterior distribution of the model parameters:

1. Draw θ_i^1 for each household using a random walk Metropolis-Hastings step with candidate density: $\mathbb{L}_i(\theta_i^1|\theta^2, \Upsilon, \Xi, Z_i)$
2. Draw θ^2 using a random walk Metropolis-Hastings step with candidate density: $\mathbb{L}\left(\theta^2|\left\{\theta_i^1\right\}, \Upsilon, \Xi, Z_i\right)$
3. Draw $\Upsilon|\left\{\theta_i^1\right\}, \Xi$ from the posterior (assuming conjugate prior)
4. Draw $\Xi|\left\{\theta_i^1\right\}$ from the posterior (assuming conjugate prior)
5. Repeat steps 1-4

To speed convergence of the algorithm, we obtain starting values from the maximum likelihood estimates, assuming homogenous parameters. We run the chain for 400,000 draws and discard the first 300,000 as burn-in. Reported estimates of the posterior distribution in the next section are based on the final 10,000 draws, thinned by retaining every 100th draw.

5.1 Computational issues

The primary computational challenge for estimation relates to calculating the expected product utility for consumers prior to their making a channel choice. This process requires repeated computation of optimal category quantities for many draws of the channel/category shocks. To overcome the computational burden, we develop an algorithm that efficiently simulates optimal category quantities. Our algorithm follows Pinjari and Bhat (2010) and is explained in detail in Appendix B. The key feature of the algorithm is that it solves for optimal quantities via a fixed number of matrix operations, rather than requiring a constrained nonlinear search. When simulating v_c and performing the Monte Carlo integration, we use 1000 draws generated from a Halton sequence.

5.2 Identification

In a non-linear model such as the one presented here, functional form and distributional assumptions inevitably contribute to parameter identification. Nevertheless, key patterns of variation in the data link unambiguously to certain parameters. We summarize these relationships in Table 4. For the distance effects in particular, we leverage quasi-experimental variation in customer-store distance that results from store entry events. This approach uses household-level changes in store distance after the inclusion of individual fixed effects and CSA time trends.

	Parameter(s)	Key identifying variation	Additional controls
Product utility (Ψ_{ck})	ψ_{ck}, $\tilde{\psi}$	category incidence	
	γ_{ck}	co-occurrence of category quantities	
	Σ_c	category quantity price response	
Retail channel utility (Ω_c)	ω_2	retail share | purchase, expected product utility	
	$\tilde{\alpha}, \tilde{\kappa}$	retail share distance response	
Budget (b)	τ, β	expenditure level, X, expenditure variance	quarter fixed effects, CSA trends
Retail fixed cost (f_2)	α, κ	retail share, expenditure level distance response	
Consideration arrival (μ)	δ, $\tilde{\beta}$	purchase incidence, $\tilde{X}$	quarter fixed effects, CSA trends
	ζ	purchase incidence distance response	

Table 4: Identification summary

Although our sampling frame falls short of providing "full coverage" in terms of observing all possible channel and category choices for each individual, it is useful to consider theoretical identification in this context. In table 4, category incidence refers to the frequency of purchases involving the category. Individual-specific category incidence rates can identify ψ_{ck}, even when categories are purchased discretely

(single category baskets) and no price variation is present. Quantity variation in multi-category baskets reflects the cross-category quantity trade-offs that identify the satiation parameters γ_{ck}. Adding price variation allows the variance of category/channel shocks to be identified. When these shock variances are freely estimated, $\frac{1}{\sqrt{\sigma_{ck}^2}}$ effectively becomes the price coefficient for category k in channel c, since utility is scaled by this variance when entering the likelihood.

In the channel choice model, the level of (additive) channel utility is normalized by setting the online utility intercept (ω_1) to zero. Average retail channel incidence rates given purchase then identify ω_2 for each consumer. The distance effect of retail entry is identified by how the retail channel incidence changes with closer proximity to a retail location after entry. Whether the effect enters here or in the fixed costs for the retail channel, f_2, is identified by whether any of the transportation costs savings are then allocated to some of the inside good categories. Note that retail fixed costs are restricted to be zero at zero distance through the specified functional from (which has flexible scale and curvature).

As with channel utility, the shopping budget and channel fixed costs also require a normalization since their levels are not separately identified. Our approach is to estimate the budget level (τ_i) and normalize the level of one channel's fixed costs. Our setting provides a natural normalization for online fixed costs, as the firm charges \$7.95 in shipping fees per order, and thus we set $f_1 = 7.95$. Finally, the consideration arrival rate intercepts δ are identified by customer-specific average purchase frequencies, while the $\tilde{\beta}$ are identified by common temporal changes in purchase rates. The effect of store distance on consideration (ζ) is identified by changes in purchase frequency with retail store distance, after controlling for individual-specific purchase rates and common temporal effects.

As a practical test of model identification, we estimated the model using simulated outcomes from in-sample data (assuming homogenous consumers) and we were able to recover the true parameter values within standard confidence intervals.

6 Results

In this section, we present our model estimates and discuss their implications. In section 6.1, we present the model estimates in raw form. In section 6.2, we derive traditional price elasticity measures for product category demand (quantities) from the raw parameter estimates. In section 6.3, we explore the effects of retail store distance on demand. Finally, we discuss the fit of the model to observed outcomes in section 6.4.

Product utility parameters				Online (1)		Retail (2)	
			het.	mean	std. dev.	mean	std. dev.
Category 1	baseline utility	ψ	Y	-1.3667	0.2463	-1.1604	0.3621
	satiation	γ	N	0.2344	0.0490	0.3375	0.0409
Category 2	baseline utility	ψ	Y	-2.3399	0.3251	-2.1109	0.3519
	satiation	γ	N	0.3640	0.0577	0.4991	0.0514
Category 3	baseline utility	ψ	Y	-2.4716	0.2903	-2.2536	0.3457
	satiation	γ	N	0.1121	0.0483	0.5115	0.0540
Category 4	baseline utility	ψ	Y	-1.8826	0.3813	-2.0685	0.3670
	satiation	γ	N	0.1879	0.0353	0.3915	0.0430
Category 5	baseline utility	ψ	Y	-2.4423	0.3673	-2.2425	0.4205
	satiation	γ	N	-0.2991	0.0464	0.1214	0.0390
Category 6	baseline utility	ψ	Y	-2.6627	0.4358	-5.6756	0.3511
	satiation	γ	N	0.2599	0.0671	0.6637	0.0975

obs = 57,554 household x quarter observations, some with multiple purchases.

Table 5: Category utility parameter estimates

6.1 Model estimates

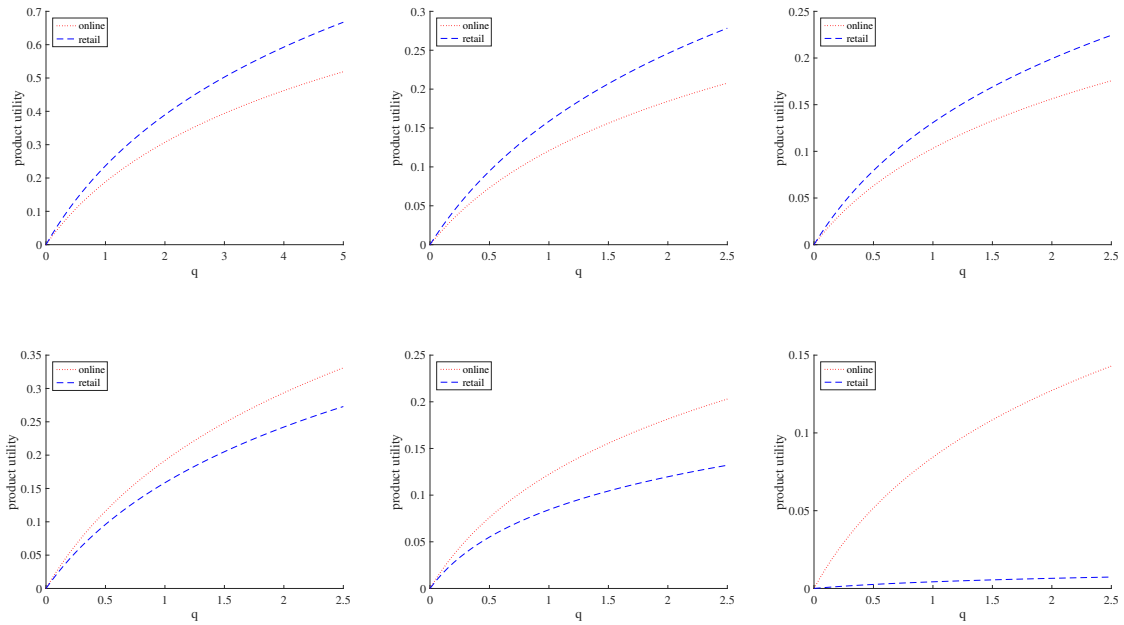
6.1.1 Product Utility

In Table 5, we present the product category utility parameter estimates. We report the mean and standard deviation of the heterogeneous baseline utilities as well as the homogeneous satiation parameters, γ. For a more intuitive representation of how these parameters relate to product utility, we also plot utility values as a function of quantity for each of the six categories by channel in Figure 10. As we would expect given the market shares, Category 1 yields the highest utility on average, and its average utility is higher in the retail channel. Categories 2 and 3 mirror category 1, but at a lower level of utility. Categories 1-3 correspond to products in which pre-purchase assessment of product fit is conventionally greater, and hence it is not surprising to see consumers derive higher utility from purchasing them in the physical store. This pattern changes for categories 4-6, in which higher utility is achieved through the online channel. For channel 6 in particular, the retail utility of purchases is very low on average, reflecting the very small expenditure share we see in observed purchases. To interpret the satiation parameter estimates, recall that larger γ values imply lower rates of satiation and hence larger average purchase quantities. In general, satiation rates are lower for the retail channel, which is consistent with the observed lower average expenditure levels in the retail channel.

In Figures 11 and 12, we plot the heterogeneity distributions of the baseline category utility parameters ψ for the online and retail channels, respectively. Most of the categories have a fatter right tail than left tail,

Figure 10: Product utility mean values as a function of purchase quantity

(a) Category 1 (b) Category 2 (c) Category 3



(d) Category 4 (e) Category 5 (f) Category 6

and some category/channel combinations have a clearly multi-modal distribution, such as tops in the retail channel. This individual-level heterogeneity in preferences for different categories will not only lead some consumers to prefer some products over others, it will also lead some consumers to prefer one channel over the other, since channel choices are determined in part by expected product utilities.

Although we allow the utility of a product category to be affected by whether the category was purchased in the last transaction (if the last transaction was in the same or previous quarter), we see limited evidence of a state dependence effect. Thus, cross-channel spillover effects are limited for the application we study. The only significant effect we find is that the utility for a category in the retail channel is lower if the consumer recently purchased that category in the retail channel.

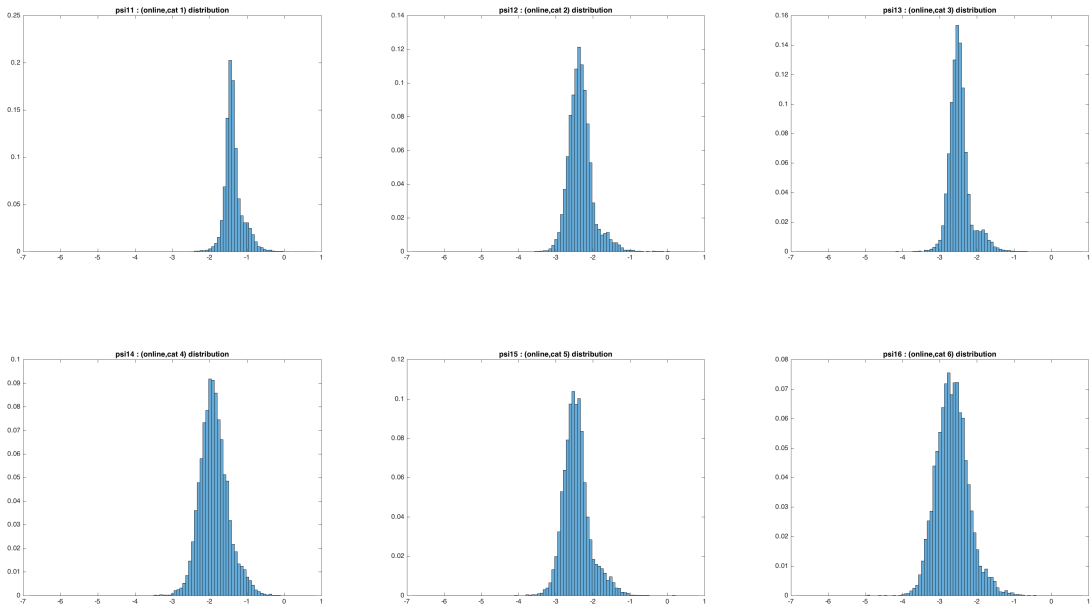
State dependence parameters (current channel X previous purchase channel)	mean	std. dev.
online X online	0.0367	0.0184
online X retail	-0.0427	0.0232
retail X retail	-0.1151	0.0101
retail X online	0.0226	0.0256

Table 6: Product category state dependence estimates ($\tilde{\psi}$)

Finally, we present the estimated covariance matrices for the product category shocks in Table 7. With

Figure 11: Online product utility intercepts, ψ_{i1k}

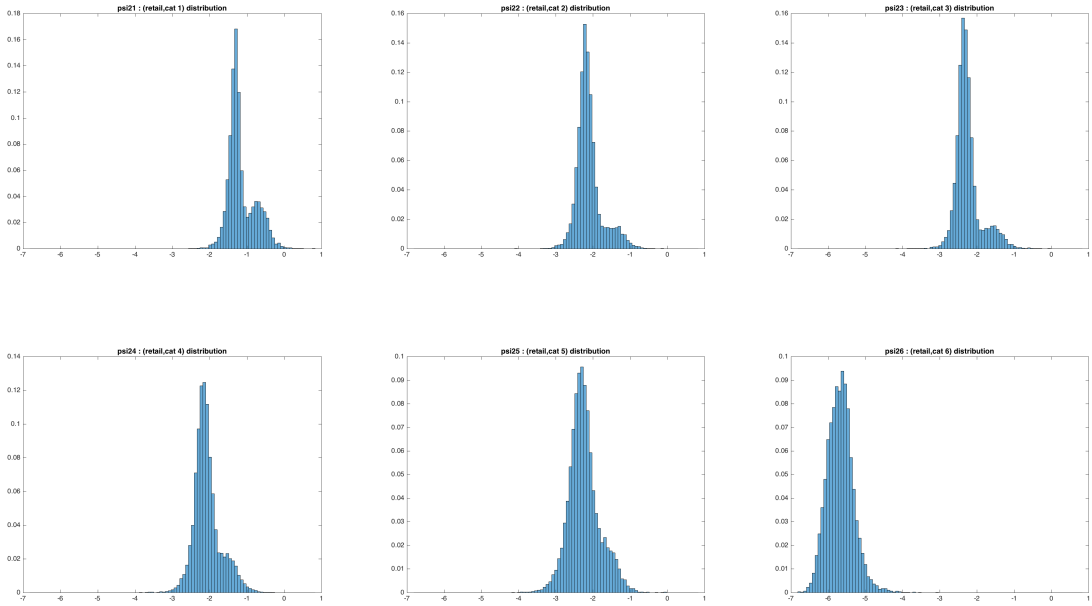
(a) Category 1 (b) Category 2 (c) Category 3



(d) Category 4 (e) Category 5 (f) Category 6

Figure 12: Retail product utility intercepts, ψ_{i2k}

(a) Category 1 (b) Category 2 (c) Category 3



(d) Category 4 (e) Category 5 (f) Category 6

the exception of category 6, shock variances are larger in the online channel, consistent with higher variability of online product fit and lower online price sensitivity. In general, we see larger shock covariances in the retail channel, suggesting that the retail channel permits more cross-category comparisons, including assessment of cross-category complementarities. Most notable is the fact that there is a covariance of 0.99 between category 3 and 6 in the retail channel. These product categories are often purchased together such that consumers choose something in category 6 to "go with" the category 3 purchase – the extent of this complementarity is understandably easier to assess with physical access to the products.

	cat. 1	cat. 2	cat. 3	cat. 4	cat. 5	cat. 6
category 1	1.6790					
category 2	0.3594	1.9931				
category 3	0.2510	0.2320	4.3258			
category 4	0.2273	0.2588	0.1701	1.5280		
category 5	0.2545	0.1538	0.1754	0.1485	2.1532	
category 6	-0.1169	0.0204	0.1180	-0.0974	0.3358	4.2101

(a) Online channel

	cat. 1	cat. 2	cat. 3	cat. 4	cat. 5	cat. 6
category 1	1.1765					
category 2	0.3875	1.5786				
category 3	0.3162	0.3300	2.4943			
category 4	0.3136	0.3315	0.4079	1.2500		
category 5	0.1351	0.0945	0.1340	0.0855	1.0814	
category 6	0.1653	0.1350	0.9868	0.3492	0.2704	5.2814

(b) Retail channel

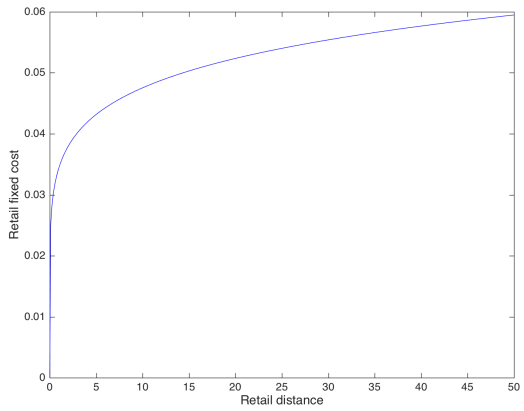
Table 7: Estimated product category shock covariances

6.1.2 Channel utility, fixed costs and consideration arrival

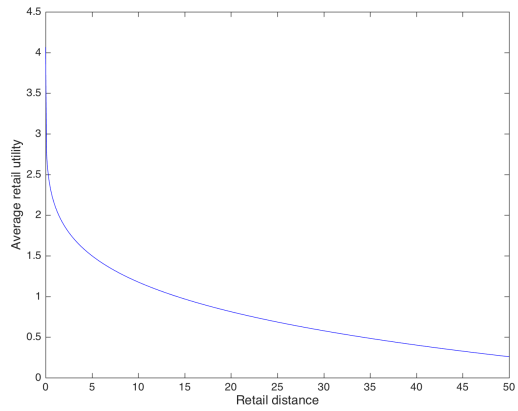
Retail utility, retail fixed cost and budget parameter estimates are shown in Table 8. The scale parameter for store distance in fixed costs for the retail channel is $\exp(-3.37) = 0.03$, which implies a 1 mile trip would cost the consumer 3 cents. As can be seen in Figure 13, transportation costs reduce the shopping budget by only six cents over a distance of 50 miles, implying that we find negligible transportation costs that are fungible with the shopping budget. By contrast, the scale parameter for store distance in retail channel utility (which is unconstrained in sign) is -1.97, which leads to a large drop in utility (3.26 units) between zero and 20 miles. We can interpret the retail utility distance effect as capturing monetary transportation costs from a different financial budget as well as time-related transportation costs.

Figure 13: Effects of store distance on retail fixed costs and utility

(a) Fixed Costs as a Function of Distance



(b) Retail Utility as a Function of Distance



		heterogeneity	mean	std dev
Retail utility parameters				
intercept	ω_2	Y	4.0643	0.6684
distance scale	$\tilde{\alpha}$	N	-1.9475	0.0806
(log) distance power	$\tilde{\kappa}$	N	-1.7633	0.0349
Retail fixed cost parameters				
(log) distance scale	α	N	-3.3666	0.0102
(log) distance power	κ	N	-1.9716	0.1209
*Expenditure parameters**				
budget intercept	τ	Y	5.6171	0.4884

obs = 57,554 household x quarter observations, some with multiple purchases.

* quarter FE and CSA time trends in β not reported for brevity.

Table 8: Retail utility, retail fixed cost and budget parameter estimates

In Table 9, we report the arrival rate coefficients. The consideration arrival parameter ζ captures the effect of retail distance on brand consideration. The negative (highly significant) coefficient for ζ implies that as distance decreases, consideration arrival rates (and hence purchase frequencies) will increase. Again to assist in interpretation, we plot the arrival rate effect multiplier $(1+d)^{\zeta}$ as a function of distance in Figure 14. Consideration events for consumers 50 miles from a retail store arrive at 81% of the rate for consumers who are located within one mile.

Consideration arrival parameters		heterogeneity	mean	std dev
intercept	δ	Y	-4.4045	0.5765
distance power	ζ	N	-0.0528	0.0094

obs = 57,554 household x quarter observations, some with multiple purchases.

* quarter FE and CSA time trends in $\tilde{\beta}$ not reported for brevity.

Table 9: Consideration arrival parameter estimates

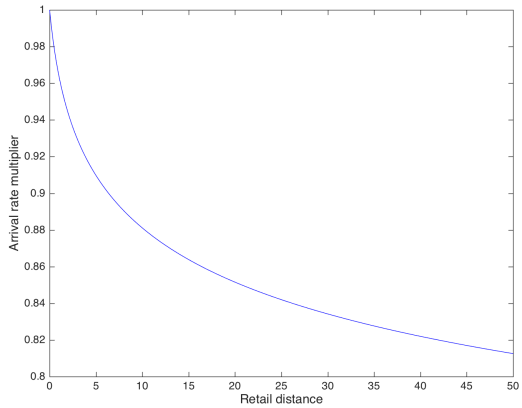


Figure 14: Effect of store distance on consideration brand arrival rate

6.1.3 Heterogeneity hyperparameters and correlations

In Table 14 of Appendix C, we report report estimates of the hyper parameters $\overline{\Upsilon}$, which capture the projection of heterogeneous parameters onto demographic variables observed by individual (age) and imputed from Census block groups (median household income, percent white population, etc.). These results suggest younger consumers prefer category 1 relative to the other categories and have smaller budgets, and that white consumers prefer the retail channel format.

Table 15 of Appendix C reports the correlation pattern among the heterogeneous parameters. In the online channel, there are strong positive correlations in baseline utility for categories 2-5. The retail channel does not exhibit the same level of correlation. Budget levels are negatively correlated with all baseline utility parameters. The retail utility intercept is negatively correlated with the online baseline category utilities for all categories, whereas it is positively correlated with retail category baseline utilities. This pattern suggests that customers who prefer shopping in the retail channel, irrespective of the product utility, also get more product utility when shopping in the retail channel. The arrival rate and budget are also positively correlated, indicating that more frequent shoppers also tend to purchase more products.

6.2 Price elasticities

Now that we have presented the full set of parameter estimates, we can turn to our main objectives of inference. In this section, we begin by deriving conventional price elasticity measures from our model estimates and data. In section 6.3, we investigate the effect of retail entry (by decreasing the distance to the nearest retailer) on firm revenues, which is of primary interest.

For unit demand price elasticities, we work with expected category/channel quantity indices, $\overline{q}_{ck} = \mu Pr(y_c = 1) E\left[q^*_{ck} | y_c = 1\right]$ and compute elasticities as $\frac{\partial \bar{q}_{ck}}{\partial p_{ck}} \frac{p_{ck}}{\bar{q}_{ck}}$. We compute these measures by quarter, and then average across individuals and quarters to produce aggregate category price elasticities. These aggregate demand elasticities with respect to price are reported in the first two columns of Table 10. The elasticities assess the change in quantity demanded if the category price is changed by one percent in the channel of interest. These measures capture cross-channel substitution and substitution to the outside good, although most substitution occurs among categories within the same channel. In all cases, price elasticities are larger in magnitude in the retail channel, which implies that markups should be lower in this channel if prices are optimized at the category level. This finding is consistent with what we observe in the data, assuming category marginal costs are constant across channels. In columns 3 and 4, we also report purchase frequency elasticities with respect to price. The small magnitude of purchase frequency elasticities indicate that the primary effect of price changes on demand operates through cross-category substitution conditional upon purchase, rather than through changes to purchase frequency.

	demand (quantity)		purchase frequency	
category	online	retail	online	retail
1	-1.559	-1.630	-0.0037	-0.0014
2	-1.850	-1.959	-0.0014	-0.0006
3	-1.525	-1.718	-0.0030	-0.0007
4	-1.550	-1.703	-0.0023	-0.0008
5	-1.429	-1.584	-0.0015	-0.0009
6	-1.523	-1.902	-0.0021	-0.0001

Table 10: Elasticities of category demand (quantity) and total expenditures with respect to price

We also calculate the elasticity of channel revenues with respect to price if the prices for all six categories experienced a uniform percentage price change. The online revenue elasticity with respect to a uniform price change for all categories in the online channel is -0.735, whereas the cross-channel elasticity for retail revenues under this change to online prices is 0.330. Conversely, for a comparable change to prices in the retail channel, the own price elasticity is -0.641, and the cross-channel elasticity for online revenues is 0.272. Thus, although own-channel revenues are inelastic with respect to price changes in that channel, there is still significant cross-channel substitution.

6.3 Store distance effects

Of particular interest are retail distance effects on total expenditures and expenditures by channel. In Table 11, we report expected quarterly channel expenditure and purchase frequency elasticities (calculated numerically) with respect to retail store distance, evaluated at the empirical distribution of customer/store locations. Our estimates imply a 10% reduction in retail store distance increases existing-customer total (both retail and online) expenditures by 0.83%. The total expenditure increase arises primarily due to an increase in the total purchase frequency (0.91%), which is attributed to increased brand consideration from the closer retail presence. In addition, decreasing the distance to the nearest store increases the probability of a retail purchase by 1.36% while the probability of online purchases declines by 2.17%. At average expenditure levels, the 10% store distance reduction translates to a 76 cent increase in retail revenues per customer per quarter and a 20 cent reduction in comparable online revenues, a net increase of 51 cents.

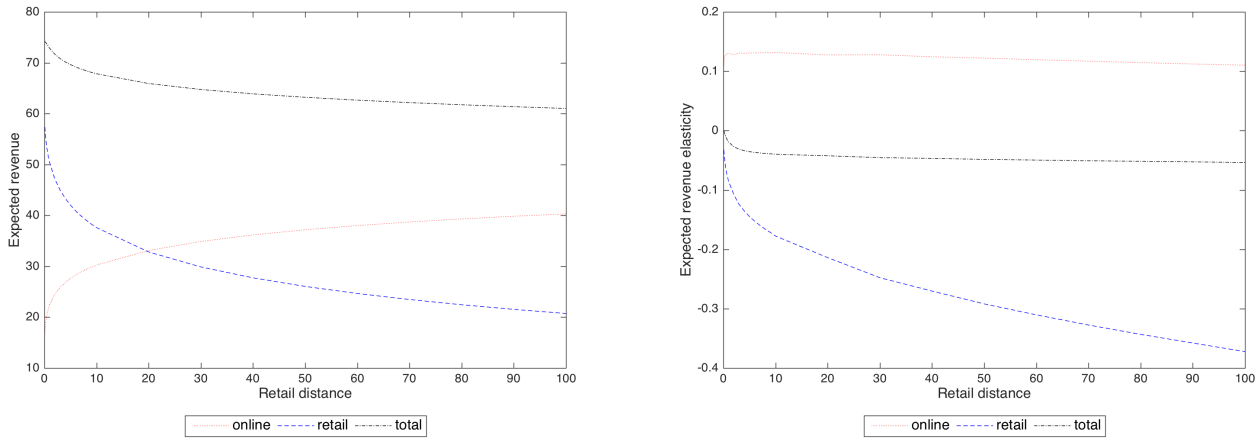
elasticities (per customer-quarter)	purchase frequency	expenditures
total	-0.091	-0.083
online	0.217	0.063
retail	-0.136	-0.215

Table 11: Elasticities with respect to retail store distance

When evaluating revenue effects, other distance distributions may be of interest. In particular, fixing retail distance to a specific value for all customers in the sample provides a measure of how the firm's customer base would respond to a uniform retail distance treatment effect (in essence, reflecting a representative consumer for the heterogeneous sample). We simulate a series of such datasets where distance treatments are varied across datasets, focusing on expected quarterly revenues by channel as a function of retail store distance. We average over individuals in the sample to obtain per capita per quarter revenue estimates. In the left hand panel of Figure 15, we plot these expected quarterly revenues by channel and total quarterly expected revenue versus retail distance. In the right hand panel of Figure 15, we plot point elasticities corresponding to the curves in the left hand panel.

In the left panel of Figure 15, total revenues increase monotonically as distance decreases, from approximately $61 per customer per quarter at 100 miles retail distance to $68 at ten miles and $74 at one mile, figures that capture the net economic benefit of shifting retail availability. It is clear that gains in total revenue track increases in retail revenue, which are only partially offset by more modest declines in online revenue. Retail and online channels contribute equally to firm revenues at a distance of approximately 20

miles from the store. In terms of elasticities (right panel of Figure 15), we see that distance elasticities for retail revenues are always negative and decline monotonically with distance. In contrast, online revenue elasticities with respect to distance are fairly constant with values between 0.11 and 0.14.



(a) Expected revenue

(b) Expected revenue elasticities

Figure 15: Expected quarterly revenue by channel with equidistant customers, as a function of retail store distance

On average we see that online revenues decrease as the distance to the nearest retail store declines from retail entry. This indicates that for the online channel, the cannibalization effect is generally larger than the consideration arrival effect. However, this is not the case for all consumers. In Figure 16, we plot the histogram of online revenue elasticities with respect to distance across consumers, in which we average the elasticity across observations for each consumer. The bulk of the distribution is to the right of the y-axis, and the mean is 0.25, substantially larger than the 0.06 value we see when averaging across observations, which suggests that consumers with less positive elasticities purchase considerably more. Furthermore, for two percent of consumers, the online revenue elasticity is negative, meaning that online expenditures increase as retail store distance decreases. For these consumers, the increase in online sales due to increased consideration is larger than the decrease due to channel switching. The considerable heterogeneity across consumers underscores the importance of tailoring retail policy to local markets and targeting retail expansion in areas that generate the highest incremental revenue.

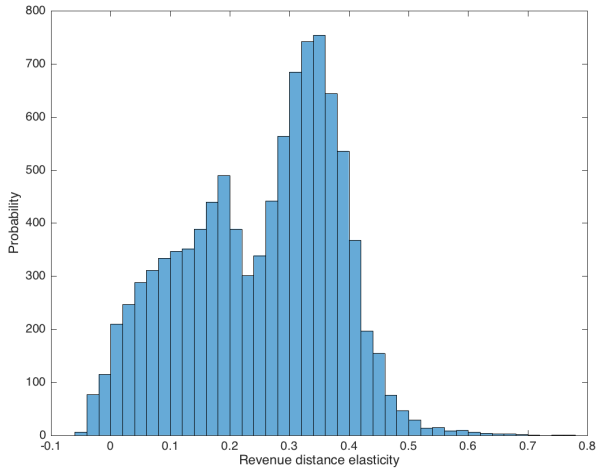


Figure 16: Distribution (across customers) of online expenditure elasticity with respect to retail store distance

6.3.1 Store distance effect decomposition

In this section, we decompose the previously reported aggregate store distance effects by isolating the impact of the various distance parameters in our model. This decomposition is useful to generate intuition about the potential impact of store entry in other environments, such as those in which the store-driven consideration effect is less important because other marketing strategies are being used to generate consumer awareness and consideration (e.g., localized advertising or pricing policies).

To perform the decomposition, we first use the parameter estimates to simulate data under the current environment. We then sequentially set each distance-related parameter to zero and re-simulate outcomes under the counterfactual. If we set the effect of store distance to zero in the budget constraint, the effect on revenues is negligible (under a penny). This again shows how consumers do not treat savings from a reduction in retail transportation costs as fungible with their shopping budget for products. By contrast, eliminating the negative distance effect in the retail utility expression (which again is how distance would enter if there was a separate budget for transportation costs) reduces online revenues by $25.22 per customer per quarter (79.5%) and increases retail revenues by $21.60 per customer per quarter (61.5%). Thus total revenues under this scenario actually decrease by $3.62 per customer per quarter (5.4%), because the prices charged in the retail channel are generally lower than in the online channel. Thus, too much store entry could undermine the firm's ability to price discriminate across channels.

If we instead were to remove the distance effect on the consideration arrival rate (in which arrivals

are reduced at larger distances), we see that online revenues would increase by \$5.96 (18.8%) and retail revenues by \$4.44 (12.6%), per customer per quarter. Total revenues increase by \$10.40 (15.6%). Thus, for its existing customers, it is the consideration effect that provides the greatest value to the firm from its entry decisions.[15]

6.4 Model fit

To assess the fit of the model, we simulate purchase occasions keeping the panel variables the same as in our estimation sample. In Appendix D, we provide distribution plots for: a) the number of simulated purchase occasions (corresponding to Figure 1 in the sample data), b) kernel density plots of purchase expenditure levels by channel (corresponding to Figure 4) and c) category average shares by channel (corresponding to Figure 5). In addition, we compare channel choice frequencies and other key moments in the simulation and data. In general, the model over-predicts the number of purchase per quarter, and thus revenues, slightly compared to that found in the data. In comparing the distribution of purchase frequency from the simulation in Figure 21 to the observed data in Figure 1, we see that the model predictions match the data closely, except for a small set high purchase-frequency consumers (≥ 10 purchases per quarter). In comparing the distributions of expenditures by channel, the data have more very small purchases than the model predictions; otherwise, the distributions are quite similar. The category shares match very well across the actual and simulated data, as does the relationship between channel choice probability and distance, shown in Figures 24 and 25.

7 Counterfactuals

In this section, we conduct counterfactual experiments to demonstrate how our results may inform channel pricing policies and retail expansion strategies.

7.1 Eliminating online shipping costs

An important issue for multi-channel firms is how to design channel-specific pricing policies. Our model can be used to assess the demand (and with cost information, profit) implications of various pricing policies.

[15] Clearly, retail entry also plays an important role in driving new customer acquisition. Revenues from store entry-induced new customer acquisition are likely significant and potentially larger than the incremental revenue from existing customers. However, new customer acquisition is beyond the scope of our model and hence we cannot quantify the relative contributions to firm revenues.

Generally, the firm can influence variable shopping costs by setting channel-specific category prices, or the firm can influence fixed shopping costs by varying online shipping fees. In our first counterfactual, we assess what would happen to firm revenues under the elimination of online shipping costs. To perform this experiment, we compare simulated outcomes under observed shipping fees ($f_1 = 7.95$) and a counterfactual scenario where the fee is eliminated ($f_1 = 0$).

We summarize the experiment in Table 12 below. For the baseline ($f_1 = 7.95$) and no shipping ($f_1 = 0$) scenarios, we report simulated quarterly average revenues, number of purchases, and average revenue per purchase for the online channel, retail channel, and both channels (total) per customer, per quarter. We then report the change in these measures going from the baseline to no shipping cost scenarios, in absolute (Δ) and percentage ($\Delta\%$) terms. Eliminating shipping fees increases total revenue by 1.17%, driven by a 0.90% increase in overall purchase frequency and a 0.99% increase in expenditures per purchase. Broken out by channel, online revenues increase by 3.88% while retail revenues decrease by 1.18%. The changes in retail channel revenues are due to switching in purchase incidence from retail to online.

Conclusively determining whether this policy is profitable for the firm requires knowledge of shipping costs and category margins. However, under fairly conservative assumptions, some back of the envelope calculations suggest it will not be. If we assume that shipping costs are the same as the shipping fee the firm charges to consumers (complete pass through), then we can evaluate the expected cost of implementing the policy and compare it to the expected incremental revenue the policy brings to the firm. If costs exceed revenues, the policy is not predicted to be profitable. If revenues exceed costs, a lower bound on average margins that makes the policy profitable could be compared to actual average margins if such information is available. Here, when the firm eliminates online shipping fees, it must pay shipping costs on all online purchases, meaning the expected cost of the policy is 0.208*7.95=$1.65 per customer per quarter. At the same time, the net revenue gain is $0.90 per customer per quarter, implying an expected net loss of at least $0.75 per customer per quarter (depending on the product markups). While fully eliminating shipping fees appears not to be profitable, there may be scope for reducing shipping fees without incurring losses. Moreover, such reductions could potentially have benefits for customer acquisition that are not accounted for here.

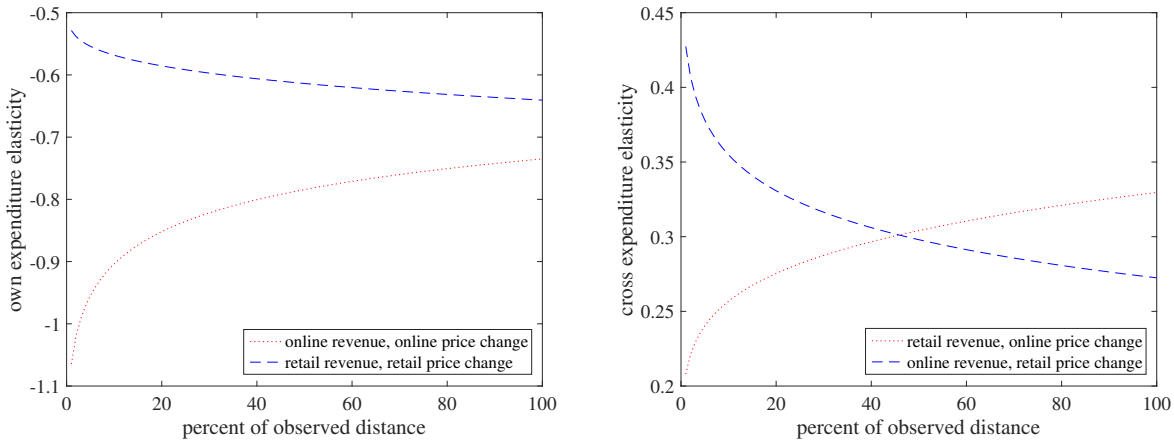
	Total (Online + Retail)			*Online*			*Retail*		
	Quarterly revenue	Quarterly purchases	Rev per purchase	Quarterly revenue	Quarterly purchases	Rev per purchase	Quarterly revenue	Quarterly purchases	Rev per purchase
baseline	77.542	0.505	140.313	35.984	0.201	160.882	41.557	0.304	131.083
no shipping	78.446	0.506	141.699	37.383	0.208	162.498	41.063	0.299	131.694
Δ	0.904	0.001	1.386	1.399	0.007	1.616	-0.494	-0.005	0.611
Δ%	1.17%	0.27%	0.99%	3.88%	3.08%	1.00%	-1.18%	-1.59%	0.46%

Table 12: No online shipping costs counterfactual results

7.2 Retail expansion

In this subsection, we demonstrate the extent to which channel-specific pricing might be undermined by retail channel expansion. First we show the effect of retail entry on own and cross-channel price elasticities, in which the prices of all categories within a channel are increased together. Second, we calculate category X channel price elasticities (in which the price of a single category within a channel is altered) in order to calculate the optimal markups for the different product categories in both channels, accounting for cross-category and cross-channel substitution. We calculate the optimal markup under two alternative assumptions: (a) total (online + retail) profits are maximized, and (b) profits are maximized by channel. We finally show the revenue implications of retail expansion, as consumers' distances to stores decline. We do this in the current environment, as well as in the hypothetical situation in which the consideration effect from retail proximity is turned off.

We start with the own and cross-channel expenditure elasticities with respect to product prices, which depend on store distance. As more retail entry occurs, higher retail channel utility will increase consumer switching from the online channel, while overall purchase frequency will increase due to higher consideration rates – the net effect of which is realized in channel expenditure levels.. In Figure 17, we graph the expenditure elasticities for the two channels as prices of all six categories within the channel are changed by the same relative amount, for distances ranging between one percent of the observed distances in the data up to the actual realizations. This exercise demonstrates what might be expected if the firm were to continue expanding its retail footprint. As expected, as consumers get closer to retail outlets, the online channel own-price elasticity increases in magnitude from -0.735 to -1.065, whereas the retail channel own-price elasticity declines in magnitude, from -0.641 to -0.529. This pattern results from cross-channel switching: consumers become more willing to switch from the online channel to the retail channel and less likely to switch in the other direction. The cross-price elasticity for retail revenues (with respect to online price changes) de-



(a) Own-channel expenditure price elasticities vs. store distance (b) Cross-channel expenditure price elasticities vs. store distance

Figure 17: Effects of store distance on channel revenue elasticities

creases from 0.330 to 0.208, whereas the cross-price elasticity for online revenues (with respect to retail price changes) increases from 0.272 to 0.427. These results have important implications for the ability of the firm to price discriminate across channels under a retail expansion strategy. Especially if the channels are managed independently, the online channel manager will have an incentive to decrease the prices of all categories with retail expansion, whereas the manager of the retail channel will have an incentive to increase prices in order to offset losses from reduced demand.

To demonstrate this further, we calculate the price elasticity for each category X channel combination in order to calculate optimal markups. When maximizing total profits over both channels, the firm's optimal category-level prices are given by:

$$\vec{p}^* = \vec{c} + (\nabla \overrightarrow{q})^{-1} \cdot \overrightarrow{q} \tag{7}$$

in which $\vec{p}^*$ is the vector of optimal prices for the twelve category/channel combinations, $\vec{c}$ are the channel X category-specific marginal costs, $\overrightarrow{q}$ are the purchase quantities, and $\nabla \overrightarrow{q}$ is the 12 X 12 matrix of quantity gradients with respect to category prices, in which the off-diagonal elements capture substitution to other categories and across the channels . If the firm maximizes its profits by channel, then the off-diagonal elements of the gradient matrix that capture substitution to categories in the other channel are set to zero (leading to a block diagonal matrix), indicating that the manager of one channel ignores the effect of its pricing decisions on the other channel's revenues. Although we do not know marginal costs, we can calculate optimal markups for both channels as a function of distance (under both assumptions regarding profit

	category	observed price ($)	maximizing total profits		maximizing channel profits	
			markup ($)	margin (%)	markup ($)	margin (%)
online channel	1	74.86	55.84	74.6	43.18	57.7
	2	69.51	58.14	83.6	38.70	55.7
	3	90.49	89.45	98.9	49.81	55.0
	4	75.14	39.72	52.9	32.16	42.8
	5	128.36	39.61	30.9	33.08	25.8
	6	105.92	110.29	104.1	76.00	71.8
retail channel	1	56.79	52.43	92.3	46.92	82.6
	2	44.58	51.62	115.8	43.71	98.05
	3	53.13	75.53	142.2	62.04	116.8
	4	31.50	31.47	99.9	28.26	89.7
	5	153.09	23.56	15.4	21.82	14.3
	6	154.72	100.56	65.0	97.11	62.8

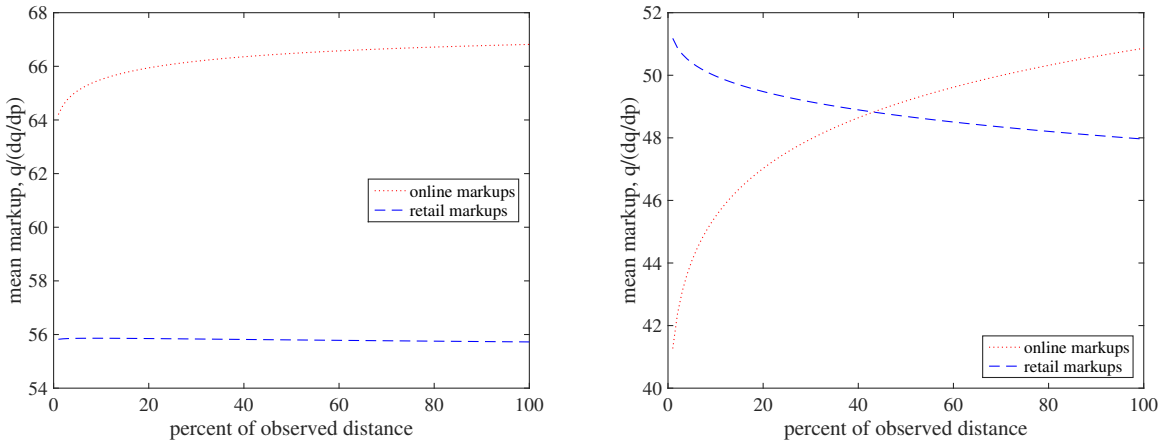
Table 13: Implied margins

maximization). We can then calculate what the costs are if we assume that the firm are currently pricing to maximize total profits or channel profits.

In Table 13, we show the observed prices in the data with the implied optimal markups under both profit maximization assumptions. If we assume the firm is maximizing total profits, then the implied marginal costs for some of the categories would be negative, since the optimal markup exceeds the observed prices, which seems unlikely. On the other hand, under the assumption of profit maximization by channel, the implied margins are more in line with what we might expect. However, since the implied category marginal costs are considerably different across channels, we are reluctant to assume that observed prices were set optimally under either strategy. Fortunately, we do not need to make either assumption to still assess the effect of entry on firm revenues, and thus the qualitative effect on profitability.

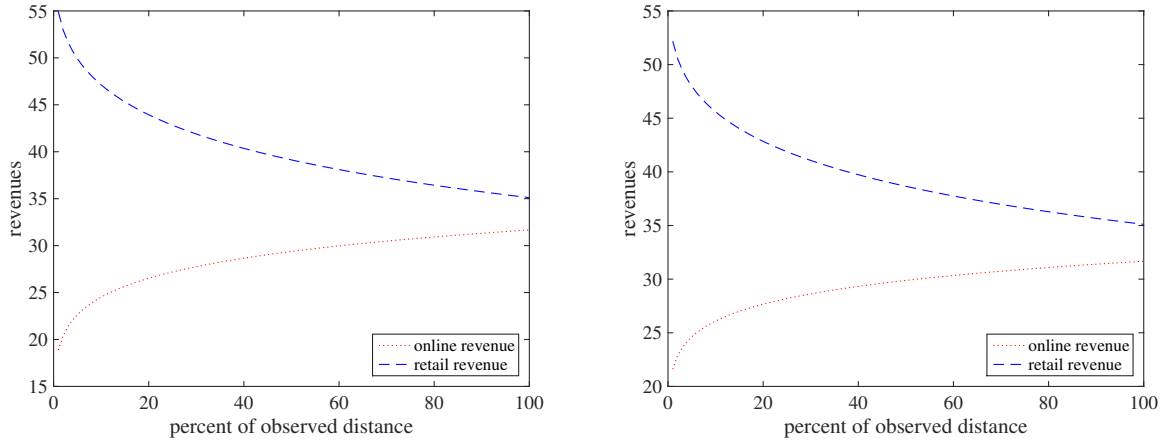
We can calculate optimal markups as distance increases from 1% of the observed distances (corresponding to a situation in which all consumers are very close to a retail store) up to the actual realizations in the data. In Figure 18, we take a simple average of the markups across categories within the channel and plot these average markups as a function of store distance. Under both profit maximization strategies, we see that optimal prices increase in the retail channel as distance shrinks (moving from right to left) and decrease in the online channel. This leads to less price differentiation, as predicted by Yoo and Lee (2011) in a scenario with vertically integrated online and retail channels.[16]

[16]Under the assumption of profit maximization by channel, price differentiation decreases only down to 42% of current distance, at which point the retail channel actually charges higher markups than the online channel. However, we find it very unlikely that the firm would be able to expand its retail footprint to the extent that the distance to the nearest store is halved for all consumers.



(a) Optimal markups vs. distance (maximizing total profits) (b) Optimal markups vs. distance (maximizing channel profits)

Figure 18: Effect of store distance on optimal markups

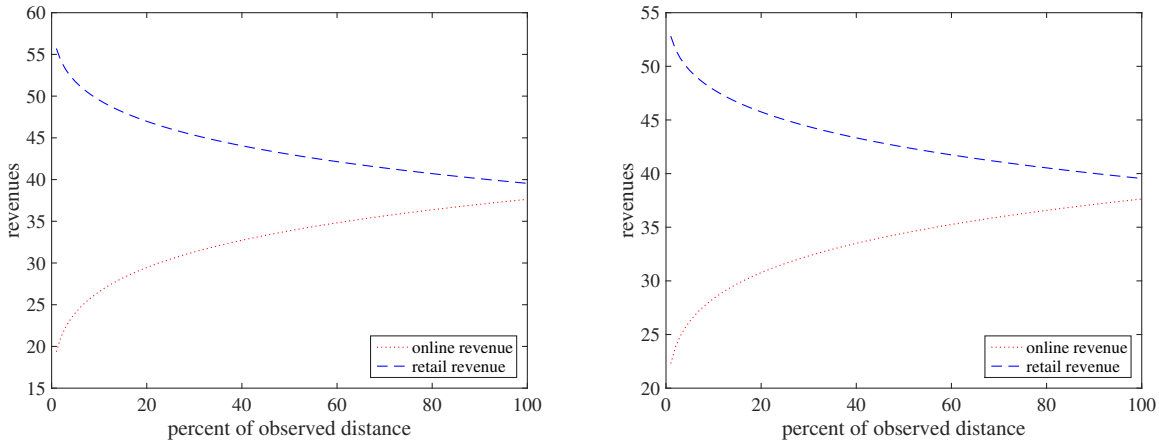


(a) Revenues vs. distance (maximizing total profits) (b) Revenues vs. distance (maximizing channel profits)

Figure 19: Effect of store distance on revenues

In Figure 19, we plot channel revenues under optimal pricing under both assumptions (total versus channel profit-maximization). Retail (and total) revenues are substantially lower when profits are maximized by channel, and online revenues are slightly higher at the observed retail distances. At smaller retail distances, retail and total revenues increase, and online revenues decrease. This result shows that the retail expansion strategy may be profitable (depending on the costs of retail expansion), despite the fact that the ability to price discriminate across channels is undermined.

If we turn off the consideration effect (effect of distance on consideration arrival rate), the retail expansion strategy is not profitable. In Figure 20, we plot channel revenues for each optimization strategy after turning off the consideration effect. Although the magnitudes of the the optimal markups are similar to those



(a) Revenues vs. distance (maximizing total profits)

(b) Revenues vs. distance (maximizing channel profits)

Figure 20: The effect of store distance on revenues, no consideration effect

when we include the consideration effect, under this counterfactual scenario, total revenues (and therefore profits) decrease as store distance declines. This follows because the lower (transportation) costs of visiting the retail channel are not recaptured by the firm through more sales, and because greater retail proximity prevents the firm from successfully price discriminating across the two channels. Given that store entry also incurs costs, the revenue decline when ignoring the effect on consideration arrival means that retail expansion would be unprofitable. The effect of entry on brand consideration is the reason that retail expansion can be justified, considering only the existing customer base.

8 Conclusion

The contributions from this research are threefold. Substantively, the paper adds to the literature that examines the demand implications of operating a mixture of online and retail channels. In addition to quantifying retail store distance effects on channel choice, we provide evidence of a promotional effect of retail proximity – i.e., retail presence raises brand consideration, leading to increased purchase frequency. We then use a structural model to estimate the size of this effect, relative to the effect of customer switching to the retail channel, which on average has lower prices in our setting. In our counterfactual exercises, we show that the effect of entry on customer switching undermines the ability of the firm to profitably price discriminate across channels.

Among existing customers, our structural estimates imply a 10% reduction in retail store distance increases total quarterly revenues by 0.82%, a result of increased brand consideration. For a small fraction of consumers (2%), we find that online revenues actually increase as a result of decreasing the distance to the nearest retail outlet – for these consumers, the consideration effect exceeds channel switching. We find no recapture of retail transportation costs in consumer purchases, given that the effect of distance on channel choice does not enter through retail channel fixed costs but instead enters retail channel utility, as it would if consumers maintained a separate budget for transportation costs.

From a methodological standpoint, we develop an integrated, utility-based model that jointly predicts purchase incidence and expenditure patterns in multiple channels and product categories. In the context of our application, this formulation allows us to draw inference on the multiple mechanisms by which channels contribute to observed patterns of demand. Our model demonstrates how to endogenize channel expenditure levels through a two-stage budgeting process, wherein first stage channel choice maximizes the expected (second stage, category quantity) utility. Additionally, our model allows for both fixed and variable transaction costs. Finally, we develop a computationally efficient algorithm to jointly estimate the multi-stage model. While our application is to semi-durables, our model and estimation methodology can be applied in any multi-channel context where the analyst has access to historical customer transaction data.

There are several potential avenues to extend the current work. One approach might compare our model formulation to those that impose alternative assumptions about the consumer's shopping decision sequence. Researchers with access to firm inventory data could investigate cross-channel assortment depth and stock-out effects. Relatedly, the role of channel interactions in the consumer's search process could be explored. While we believe that our focus on seasonal semi-durables limits the scope for dynamics in our analysis to some extent (e.g., the ability to learn is restricted by the fact that products within the observed categories are perpetually changing), accounting for these aspects would be an interesting and challenging direction for future work.

References

Ansari, A., C. Mela, and S. Neslin (2008). Customer channel migration. *Journal of Marketing Research 45*(1), 60–76.

Avery, J., T. J. Steenburgh, J. Deighton, and M. Caravella (2012). Adding bricks to clicks: Predicting the patterns of cross-channel elasticities over time. *Journal of Marketing 76*(3), 96–111.

Bell, D. R., S. Gallino, and A. Moreno (2017). Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science*.

Bell, D. R., T.-H. Ho, and C. S. Tang (1998). Determining where to shop: Fixed and variable costs of shopping. *Journal of Marketing Research 35*(3), 352–369.

Bell, D. R. and J. M. Lattin (1998). Shopping behavior and consumer preference for store price format: Why large basket shoppers prefer EDLP. *Marketing Science 17*(1), 66–88.

Bhat, C. R. (2008). The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological 42*(3), 274–303.

Biyalogorsky, E. and P. Naik (2003). Clicks and mortar: The effect of on-line activities on off-line sales. *Marketing Letters 14*(1), 21–32.

Bronnenberg, B. J. (2018). Retailing and consumer demand for convenience. In *Handbook of Research on Retailing*, pp. 17. Edward Elgar Publishing.

Chevalier, J. A., A. K. Kashyap, and P. E. Rossi (2003). Why don't prices rise during periods of peak demand? Evidence from scanner data. *American Economic Review 93*(1), 15–37.

Deleersnyder, B., I. Geyskens, K. Gielens, and M. Dekimpe (2002). How cannibalistic is the internet channel? A study of the newspaper industry in the United Kingdom and the Netherlands. *International Journal of Research in Marketing 19*(4), 337–348.

Dzyabura, D., S. Jagabathula, and E. Muller (2019). Accounting for discrepancies between online and offline product evaluations. *Marketing Science 38*(1), 88–106.

Forman, C., A. Ghose, and A. Goldfarb (2009). Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management Science 55*(1), 47–57.

Gentzkow, M. (2007). Valuing new goods in a model with complementarity: Online newspapers. *American Economic Review 97*(3), 713–744.

Geyskens, I., K. Gielens, and M. Dekimpe (2002). The market valuation of internet channel additions. *Journal of Marketing 66*(2), 102–119.

Gordon, B. R., A. Goldfarb, and Y. Li (2013). Does price elasticity vary with economic growth? A cross-category analysis. *Journal of Marketing Research 50*(1), 4–23.

Huang, Y. and B. J. Bronnenberg (2018). Pennies for your thoughts: Costly product consideration and purchase quantity thresholds. *Marketing Science 37*(6), 1009–1028.

Kim, J., G. M. Allenby, and P. E. Rossi (2002). Modeling consumer demand for variety. *Marketing Science 21*(3), 229–250.

Kim, Y., R. Telang, W. B. Vogt, and R. Krishnan (2010). An empirical analysis of mobile voice service and sms: A structural model. *Management Science 56*(2), 234–252.

Kireyev, P., V. Kumar, and E. Ofek (2017). Match your own price? Self-matching as a retailer's multichannel pricing strategy. *Marketing Science 36*(6), 908–930.

Lemon, K. N. and P. C. Verhoef (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing 80*(6), 69–96.

Li, H. and P. K. Kannan (2014). Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research 51*(1), 40–56.

Neslin, S. A., D. Grewal, R. Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, and P. C. Verhoef (2006). Challenges and opportunities in multichannel customer management. *Journal of Service Research 9*(2), 95–112.

Neslin, S. A. and V. Shankar (2009). Key issues in multichannel customer management: Current knowledge and future directions. *Journal of interactive marketing 23*(1), 70–81.

Pauwels, K. and S. A. Neslin (2015). Building with bricks and mortar: The revenue impact of opening physical stores in a multichannel environment. *Journal of Retailing 91*(2), 182–197.

Pinjari, A. R. and C. Bhat (2010). An efficient forecasting procedure for Kuhn-Tucker consumer demand model systems. Technical report, University of Texas.

Soysal, G. and A. Zentner (2014). Measuring e-commerce concentration effects in the apparel industry. Technical report, University of Texas at Dallas.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science 4*(3), 199–214.

Thomassen, Ø., H. Smith, S. Seiler, and P. Schiraldi (2017). Multi-category competition and market power: A model of supermarket pricing. *American Economic Review 107*(8), 2308–51.

Van Nierop, J., P. Leeflang, M. Teerling, and K. Huizingh (2011). The impact of the introduction and use of an informational website on offline customer buying behavior. *International Journal of Research in Marketing 28*(2), 155–165.

Venkatesh, R. and R. Chatterjee (2006). Bundling, unbundling, and pricing of multiform products: The case of magazine content. *Journal of Interactive Marketing 20*(2), 21–40.

Verhoef, P. C., P. K. Kannan, and J. J. Inman (2015). From multi-channel retailing to omni-channel retailing: Introduction to the special issue on multi-channel retailing. *Journal of Retailing 91*(2), 174–181.

Wang, K. and A. Goldfarb (2017). Can offline stores drive online sales? *Journal of Marketing Research 54*(5), 706–719.

Yoganarasimhan, H. (2012). Cloak or flaunt? the fashion dilemma. *Marketing Science 31*(1), 74–95.

Yoganarasimhan, H. (2017). Identifying the presence and cause of fashion cycles in data. *Journal of Marketing Research 54*(1), 5–26.

Yoo, W. S. and E. Lee (2011). Internet channel entry: A strategic analysis of mixed channel structures. *Marketing Science 30*(1), 29–41.

Zettelmeyer, F. (2000). Expanding to the internet: Pricing and communication strategies when firms compete on multiple channels. *Journal of Marketing Research 37*(3), 292–308.

Appendices

A Optimality conditions

A.1 First stage optimality

In the first stage decision, the customer determines which channel to visit. This decision depends on the expected utility from products purchased through the channel (v_c), the deterministic utility obtained from simply visiting the channel ($\bar{\Omega}_c$), and the channel utility shock (ξ_c). The consumer chooses to visit channel c if the expected utility is higher than for visiting the other channel and higher than for allocating the entire budget to the outside good. The optimality conditions may thus be written:

$$y_c = 1 \Longrightarrow \tag{8a}$$

$$v_c + \bar{\Omega}_c + \xi_c > v_j + \bar{\Omega}_j + \xi_j \quad \forall j \neq c \tag{8b}$$

$$v_c + \bar{\Omega}_c + \xi_c > \log(b) \tag{8c}$$

The deterministic utility from channel visitation ($\bar{\Omega}_c$) can be evaluated analytically (and hence inexpensively) from model parameters and data. However, no analytical expressions are available for the expected utility from product purchases (v_c). Instead, we rely on simulation methods to numerically approximate the expected product utility. The simulation-based approach in turn requires that we be able to solve for optimal category quantities ($\vec{q}_c{}^*$) given draws of the category shocks ($\vec{\varepsilon}_c^*$). In Appendix B, we develop a highly efficient numerical method to compute optimal category quantities and expected product utilities.

A.2 Second stage optimality

For the second-stage decision (allocation of the budget across categories and the outside good), both the objective (utility) function and the constraint equation are continuously differentiable in the control variables ($\overrightarrow{q}$), so the Kuhn-Tucker theorem may be applied to solve the model. The second stage is conditioned upon channel choice, $y_c = 1$, which implies channel fixed costs are sunk at the second decision stage.

The Lagrangian for the problem is: $\mathscr{L} = U_c(\vec{q}_c, z) - \lambda \left(\left(\sum_{k=1}^{K} q_{ck} p_{ck} \right) (1 + r_c) + f_c + z - b \right)$

The optimal quantities then satisfy the Kuhn-Tucker (KT) conditions:

$\frac{\partial \mathscr{L}}{\partial z} = 0$, $\frac{\partial \mathscr{L}}{\partial q_{ck}} q_{ck} = 0$, $\frac{\partial \mathscr{L}}{\partial q_{ck}} \leq 0$, $q_{ck} \geq 0$, $z > 0$ for all $k = 1, \ldots, K$

The KT conditions for the consumer's problem reduce to:

$$
\begin{aligned}
\frac{\partial \mathscr{L}}{\partial z} =& \frac{1}{z} - \lambda = 0 \Longrightarrow \lambda = \frac{1}{z} \\
\frac{\partial \mathscr{L}}{\partial q_{ck}} =& \Psi_{ck}\,\gamma_{ck}\,\frac{1}{\frac{q_{ck}}{\gamma_{ck}}+1}\frac{1}{\gamma_{ck}} - \lambda\, p_{ck}\,(1+r_c) \leq 0 \\
\Longrightarrow& \frac{\Psi_{ck}}{\frac{q_{ck}}{\gamma_{ck}}+1} - \frac{p_{ck}\,(1+r_c)}{b - f_c - \left(\sum\limits_{k=1}^{K} q_{ck}p_{ck}\right)(1+r_c)} \leq 0 \\
\Longrightarrow& \varepsilon_{ck} \leq \log\left[\frac{p_{ck}\,(1+r_c)\left(\frac{q_{ck}}{\gamma_{ck}}+1\right)}{b - f_c - \left(\sum\limits_{k=1}^{K} q_{ck}p_{ck}\right)(1+r_c)}\right] - \bar{\Psi}_{ck} \equiv g_{ck}\,(q_{ck})
\end{aligned}
\tag{9a}
$$

where we define the $g\,(\cdot)$ function for notational convenience. The equality condition $\varepsilon_{ck} = g_{ck}(q_{ck})$ holds for chosen categories, whereas the inequality condition $\varepsilon_{ck} < g_{ck}(q_{ck})$ holds for categories that are not chosen.

A.3 Likelihood of product purchase conditional on arrival

The Kuhn-Tucker conditions are used to form the likelihood for the observed quantity vector when a purchase is observed. Without loss of generality, let the first M ($M \geq 1$) categories be chosen. We partition the channel/category shocks $\vec{\varepsilon}_c$ in two parts, with the $Mx1$ vector $\vec{\varepsilon}_{cM}$ corresponding to chosen alternatives and the $(K-M)x1$ vector $\vec{\varepsilon}_{c\widetilde{M}}$ corresponding to non-chosen alternatives. With this convention, the probability of observing quantities $\vec{q}_c$ may be written:

$$
\begin{aligned}
\ell(y_c = 1, \vec{q}_c = \vec{q}_c^{\,*}) =& \int I\,[y_c = 1]\,I\,[\vec{q}_c = \vec{q}_c^{\,*}]\,dF(\vec{\xi})dF(\vec{\varepsilon}) \\
=& \int I\left[v_c + \bar{\Omega}_c + \xi_c > \log(b)\right] \prod_{j \neq c} I\left[v_c + \bar{\Omega}_c + \xi_c > v_j + \bar{\Omega}_j + \xi_j\right] dF(\vec{\xi}) \\
&\times \left|J_{\vec{\varepsilon}_{cM} \to \vec{q}^{\,*}_{cM}}\right| \int \prod_{i=1}^{M} I\,(\varepsilon_{ci} = g_{ci}) \prod_{i=M+1}^{K} I\,(\varepsilon_{ci} < g_{ci})\,dF(\vec{\varepsilon}) \\
=& \int_{\xi_c = log(b) - v_c - \bar{\Omega}_c}^{\infty} \int_{\xi_j = -\infty}^{v_c + \bar{\Omega}_c + \xi_c - v_j - \bar{\Omega}_j} h\left(v_c + \bar{\Omega}_c\right) h\left(v_j + \bar{\Omega}_j\right) d\xi_c d\xi_j \\
&\times \left|J_{\vec{\varepsilon}_{cM} \to \vec{q}^{\,*}_{cM}}\right| \int_{\varepsilon_{c,M+1} = -\infty}^{g_{c,M+1}} \cdots \int_{\varepsilon_{c,K} = -\infty}^{g_{c,K}} \phi\left(\vec{\varepsilon}_{c\widetilde{M}} | \vec{\varepsilon}_{cM}\right) d\vec{\varepsilon}_{c\widetilde{M}} \\
=& \frac{exp\left(v_c + \bar{\Omega}_c\right)}{\sum_{j=1}^{C} exp\left(v_j + \bar{\Omega}_j\right)} \left(1 - \prod_{j=1}^{C} H\left(log\,(b) - v_j - \bar{\Omega}_j\right)\right) \\
&\times \left|J_{\vec{\varepsilon}_{cM} \to \vec{q}^{\,*}_{cM}}\right| \int_{\varepsilon_{c,M+1} = -\infty}^{g_{c,M+1}} \cdots \int_{\varepsilon_{c,K} = -\infty}^{g_{c,K}} \phi\left(\vec{\varepsilon}_{c\widetilde{M}} | \vec{\varepsilon}_{cM}\right) d\vec{\varepsilon}_{c\widetilde{M}}
\end{aligned}
\tag{10}
$$

in which $h(\cdot)$ and $H(\cdot)$ are the pdf and cdf of the standard extreme value (Gumbel) distribution ($var[\xi_c] = 1$), and $\phi(\cdot)$ is the normal distribution pdf. In general, there is no closed form for truncated integration over $\vec{\varepsilon}_{c\widetilde{M}}$, which follows a multivariate normal (MVN) distribution. However, MVN integration over a rectangular region (as in the case of the ε shocks) can be approximated via GHK simulation.

Next we derive a closed form expression for the Jacobian determinant. Elements of the Jacobian are given by:

$$\begin{aligned}
J_{ij} =& \frac{\partial \varepsilon_{ci}}{\partial q_{cj}} = \frac{\partial g_{ci}(q_{ci})}{\partial q_{cj}} \quad \forall i, j = 1, ..., M \\
=& \frac{1}{(q_{ci} + \gamma_{ci})} I(i = j) + \frac{p_{cj}(1 + r_c)}{b - f_c - \left(\sum_{k=1}^{K} q_{ck} p_{ck}\right)(1 + r_c)} \\
=& \frac{1}{(q_{ci} + \gamma_{ci})} I(i = j) + \frac{p_{cj}(1 + r_c)}{z}
\end{aligned}$$

The Jacobian determinant is then given by:

$$\begin{aligned}
\left| J_{\vec{\varepsilon}_{cM} \to \vec{q}^*_{cM}} \right| =& \begin{vmatrix} \frac{1}{q^*_{c1}+\gamma_{c1}} + \frac{p_{c1}(1+r_c)}{z} & \frac{p_{c2}(1+r_c)}{z} & \cdots & \frac{p_{cM}(1+r_c)}{z} \\ \frac{p_{c1}(1+r_c)}{z} & \frac{1}{q^*_{c2}+\gamma_{c2}} + \frac{p_{c2}(1+r_c)}{z} & \cdots & \frac{p_{cM}(1+r_c)}{z} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p_{c1}(1+r_c)}{z} & \frac{p_{c2}(1+r_c)}{z} & \cdots & \frac{1}{q^*_{cM}+\gamma_{cM}} + \frac{p_{cM}(1+r_c)}{z} \end{vmatrix} \\
=& \frac{1}{\prod_{i=1}^{M}(q_{ci} + \gamma_{ci})} \left[1 + \frac{(1 + r_c)}{z} \sum_{i=1}^{M} p_{ci}(q_{ci} + \gamma_{ci}) \right] \\
=& \frac{1}{\prod_{i=1}^{M}(q_{ci} + \gamma_{ci})} \left[1 + \frac{(1 + r_c)}{b - f_c - \left(\sum_{k=1}^{K} q_{ck} p_{ck}\right)(1 + r_c)} \sum_{i=1}^{M} p_{ci}(q_{ci} + \gamma_{ci}) \right]
\end{aligned}$$

We close this section by noting that in equation 10 above, the expected product utility obtained from visiting channel , v_c, must be computed numerically, as described in Section B below. To make estimation computationally tractable, we devise a highly efficient numerical algorithm to compute the expected product utilities.

A.4 Unconditional likelihood

Equation 10 is the likelihood of observing the channel and category outcomes conditional upon observing a transaction. We also need to account for the likelihood of not observing a transaction when a consideration event occurs. A consideration event can result in no-purchase if the consumer decides to allocate the entire

budget to the outside good due to: a) sufficiently negative channel shocks (ξ_c) at the channel choice stage, or b) sufficiently negative category shocks (ε_{ck}) realized after visiting a channel. Both possibilities are included in the following likelihood expression for no-purchase given a consideration event:

$$\begin{aligned}
\ell^0 \equiv \ell(\sum_{c=1}^{C} y_c = 0 \cup \sum_{c=1}^{C} \vec{q}_c = \vec{0} \,|\, y_c = 1) = & \int \prod_{c=1}^{C} I\left[v_c + \bar{\Omega}_c + \xi_c < \log(b)\right] dF(\xi) \\
& + \sum_{c=1}^{C} \int I\left[v_c + \bar{\Omega}_c + \xi_c > \log(b)\right] \prod_{j \neq c} I\left[v_c + \bar{\Omega}_c + \xi_c > v_j + \bar{\Omega}_j + \xi_j\right] dF(\xi) \\
& \times \int \prod_{i=1}^{K} I\left(\varepsilon_{ci} < g_{ci}(0) = \log\left[\frac{p_{ck}(1+r_c)}{b - f_c}\right] - \bar{\Psi}_{ck}\right) dF(\varepsilon) \\
= & \prod_{j=1}^{C} H\left(log(b) - v_j - \bar{\Omega}_j\right) \\
& + \sum_{c=1}^{C} \frac{exp\left(v_c + \bar{\Omega}_c\right)}{\sum_{j=1}^{C} exp\left(v_j + \bar{\Omega}_j\right)} \left(1 - \prod_{j=1}^{C} H\left(log(b) - v_j - \bar{\Omega}_j\right)\right) \\
& \times \int \prod_{i=1}^{K} I\left(\varepsilon_{ci} < g_{ci}(0) = \log\left[\frac{p_{ck}(1+r_c)}{b - f_c}\right] - \bar{\Psi}_{ck}\right) dF(\varepsilon)
\end{aligned} \tag{11}$$

To form the unconditional likelihood, we need to integrate over the (Poisson) distribution of unobserved consideration events that can rationalize the set of observed purchase occasions. Suppose for example that the Poisson arrival rate of consideration events is μ and that we observe L purchase events. One possibility is that L consideration events occurred and a purchase happened at each consideration event. Alternatively, there could have been $L+1$ consideration events with L purchases and 1 no-purchase, $L+2$ consideration events with L purchases and 2 no-purchases, and so on. In general, given $L+k$ consideration events, the probability of L purchases follows a binomial distribution with "success" probability $1-\ell^0$. Therefore, the number of observed purchases is Poisson distributed, with rate parameter $\mu\left(1-\ell^0\right)$, as shown below:

$$\ell(L) = \sum_{k=0}^{\infty} Poisson(L+k, \mu) \begin{pmatrix} L+k \\ k \end{pmatrix} \left[\ell^0\right]^k \left[1-\ell^0\right]^L \tag{12}$$

$$= \frac{e^{-\mu}}{L!} \left[\mu\left(1-\ell^0\right)\right]^L \sum_{k=0}^{\infty} \frac{\left[\mu\ell^0\right]^k}{k!} = \frac{e^{-\mu\left(1-\ell^0\right)}}{L!} \left[\mu\left(1-\ell^0\right)\right]^L \tag{13}$$

The structure of our model is such that the arrival rate of consideration events (μ) can change every quarter. However, because the likelihood of no purchase conditional on arrival (ℓ^0) changes after every transaction (due to potential state dependence) and because we never observe more than one transaction per day, it is convenient to formulate the likelihood at the daily level. We denote the *daily* consideration arrival rate using $\mu_t = \mu/T_t$,where T_t is the number of days in quarter t and μ is the quarterly consideration arrival rate. The

likelihood of observing zero purchases on a given day in quarter t is then:

$$\ell(L=0) = e^{-\mu_t\left(1-\ell^0\right)}, \tag{14}$$

whereas the probability of observing one purchase on a given day in quarter t is:

$$\ell(L=1) = \mu_t\left(1-\ell^0\right)e^{-\mu_t\left(1-\ell^0\right)}. \tag{15}$$

Thus the unconditional likelihood for customer i on day d of quarter t (with or without a transaction) is:

$$\mathbb{L}_{itd} = e^{-\mu_{it}\left(1-\ell^0_{itd}\right)}\left[I(\sum_c y_{itdc}=0) + \mu_{it}\left(1-\ell^0\right)\sum_c I(y_{itdc}=1)\ell(y_{itdc}=1,\vec{q}_{itdc}=\vec{q}^*_{itdc})\right] \tag{16}$$

where the conditional product purchase likelihood, $\ell(y_{itdc}=1,\vec{q}_{itdc}=\vec{q}^*_{itdc})$, is given in equation 10.

A.5 Joint Likelihood

We can write the likelihood for the observed purchase history of customer i in quarter t as:

$$\mathbb{L}_{it} = \prod_{d=1}^{T_t}\mathbb{L}_{itd} \tag{17}$$

The likelihood of all observations for customer i and the total likelihood as a function of all model parameters (Θ) are then given by:

$$\mathbb{L}_i = \prod_{t=1}^{T}\mathbb{L}_{it} \quad , \quad \mathbb{L}(\Theta) = \prod_{i=1}^{I}\mathbb{L}_i \tag{18}$$

where T is the number of quarters in our sample and I is the number of customers in our sample.

B Computation of optimal quantities and expected utilities

We first demonstrate the calculation of optimal quantities when the set of chosen categories is known, and we then describe the procedure to determine the set of chosen categories. For a given draw of the category shocks (ε), without loss of generality assume the first M categories are chosen. For chosen (inside good) categories, the Kuhn-Tucker condition $\frac{\partial\mathscr{L}}{\partial q_{ck}}=0$ implies the optimal quantity is given by $q^*_{ck} = \gamma_{ck}\left[\frac{\Psi_{ck}}{p_{ck}(1+r_c)\lambda}-1\right]$. Similarly, the Kuhn-Tucker condition $\frac{\partial\mathscr{L}}{\partial z}=0$ implies the optimal outside good expenditure is given by $z^*=\frac{1}{\lambda}$. Next, we use the budget constraint equation to eliminate the Lagrange multiplier λ: $b = f_c + z^* + (1+r_c)\sum_{k=1}^{M} q^*_{ck}p_{ck}$. Substituting for q^*_{ck} and z^* and solving for λ gives $\lambda = \frac{1+\sum_{k=1}^{M}\gamma_{ck}\Psi_{ck}}{b-f_c+(1+r_c)\sum_{k=1}^{M}\gamma_{ck}p_{ck}}$.

Then, substituting this expression for λ back into the optimal quantity equations above gives:

$$q_{ck}^{*} = \gamma_{ck}\left(\frac{\Psi_{ck}\left(b - f_c + (1+r_c)\sum_{k=1}^{M}\gamma_{ck}p_{ck}\right)}{p_{ck}(1+r_c)\left(1+\sum_{k=1}^{M}\gamma_{ck}\Psi_{ck}\right)} - 1\right)$$

and

$$z^{*} = \frac{b - f_c + (1+r_c)\sum_{k=1}^{M}\gamma_{ck}p_{ck}}{1+\sum_{k=1}^{M}\gamma_{ck}\Psi_{ck}}$$

To find the optimal set of chosen categories, we use the "enumerative" algorithm of Pinjari and Bhat (2010). The method is based on the insight that the price normalized baseline utilities $\left(\frac{\Psi_{ck}}{p_{ck}}\right)$ of chosen (non-chosen) goods are greater than or equal to (less than) the Lagrange multiplier, which is an implicit function of the set of chosen categories. The algorithm to predict quantity choices thus proceeds as follows:

1. Take a draw of the product shocks (ε)
2. Compute the price normalized baseline utilities for all K categories: $\Upsilon_k = \frac{\Psi_{ck}}{p_{ck}}$
3. Sort the categories from highest to lowest Υ; denote quantities sorted in this order with a tilde, e.g. $\tilde{\Upsilon}$
4. Iteratively compute the Lagrange multiplier, assuming the first m categories are chosen: $\lambda^{m} = \frac{1+\sum_{k=1}^{m}\tilde{\gamma}_{ck}\tilde{\Psi}_{ck}}{b-f_c+(1+r_c)\sum_{k=1}^{m}\tilde{\gamma}_{ck}\tilde{p}_{ck}}$
5. Determine the number of chosen categories by the relation: $M = \sum_{m=1}^{K} I\left(\tilde{\Upsilon}_m \geq \lambda^{m}\right)$
6. Compute the optimal quantities for the chosen (sorted) categories as $\tilde{q}_{cj}^{*} = \tilde{\gamma}_{cj}\left(\frac{\tilde{\Psi}_{cj}\left(b-f_c+(1+r_c)\sum_{k=1}^{M}\tilde{\gamma}_{ck}\tilde{p}_{ck}\right)}{\tilde{p}_{cj}(1+r_c)\left(1+\sum_{k=1}^{M}\tilde{\gamma}_{ck}\tilde{\Psi}_{ck}\right)} - 1\right)$ for $j \leq M$ and $\tilde{q}_{ck}^{*} = 0$ for $j > M$
7. Invert the sort order to restore the original category ordering, yielding q_{ck}^{*}

Note that the optimal outside good expenditure z^{*} may be computed in step 6 using $z^{*} = \frac{b-f_c+(1+r_c)\sum_{k=1}^{M}\tilde{\gamma}_{ck}\tilde{p}_{ck}}{1+\sum_{k=1}^{M}\tilde{\gamma}_{ck}\tilde{\Psi}_{ck}}$ or in step 7 using $z^{*} = b - f_c - (1+r_c)\sum_{k=1}^{K} q_{ck}^{*}p_{ck}$. This procedure may be vectorized so that solutions may be sought for simultaneously for all the draws (across all observations), yielding a highly efficient polynomial time algorithm (proportional to $N \cdot D \cdot K^2$ operations, where N, D and K are respectively the number of observations, draws and categories). This "enumerative" algorithm is several orders of magnitude faster that

the leading alternative procedure, which would require using a (nested) nonlinear optimization procedure to solve for optimal category quantities, for each draw of the product shocks (ε).

Once optimal quantities are in hand, it is a simple matter to evaluate the expected (indirect) utility by substituting optimal quantities into equation (5d) and averaging across draws of the product shocks.

C Heterogeneity correlations and hyperparameters

Hyper-parameters ($\overline{\Upsilon}$)			Intercept	Age	White	College	HH Income	Rural	Commute
Product utility parameters - online									
Category 1	intercept	ψ_{11}	-1.2398***	-0.0033	0.1186	-0.2103	0.0007	0.1740	-0.0023
			(0.2108)	(0.0024)	(0.1609)	(0.1580)	(0.0007)	(0.1066)	(0.0041)
Category 2	intercept	ψ_{12}	-1.5507***	-0.0140***	-0.0197	0.1644	-0.0021	0.0956	-0.0045
			(0.2263)	(0.0026)	(0.1680)	(0.1652)	(0.0008)	(0.1001)	(0.0043)
Category 3	intercept	ψ_{13}	-2.0493***	-0.0092**	-0.0838	0.1407	-0.0009	0.1622	-0.0002
			(0.2218)	(0.0028)	(0.1829)	(0.1911)	(0.0009)	(0.1166)	(0.0046)
Category 4	intercept	ψ_{14}	-0.8232***	-0.0203**	-0.0359	0.0217	-0.0011	0.1214	-0.0045
			(0.1808)	(0.0023)	(0.1445)	(0.1440)	(0.0008)	(0.0880)	(0.0041)
Category 5	intercept	ψ_{15}	-1.5071***	-0.0128**	-0.0670	-0.2762	-0.0016	0.0946	-0.0025
			(0.1878)	(0.0024)	(0.1355)	(0.1602)	(0.0007)	(0.1095)	(0.0039)
Category 6	intercept	ψ_{16}	-0.9244**	-0.0207***	-0.2958	-0.4499	-0.0017	-0.2580	-0.0079
			(0.2806)	(0.0032)	(0.2221)	(0.1949)	(0.0009)	(0.0992)	(0.0048)
Product utility parameters - retail									
Category 1	intercept	ψ_{21}	-1.2292***	-0.0002	0.0928	-0.0515	0.0010	-0.0781	-0.0020
			(0.1622)	(0.0021)	(0.1346)	(0.1307)	(0.0005)	(0.0921)	(0.0033)
Category 2	intercept	ψ_{22}	-1.7413***	-0.0102***	0.1309	0.0282	0.0001	0.0503	-0.0031
			(0.1936)	(0.0021)	(0.1601)	(0.1418)	(0.0006)	(0.0921)	(0.0038)
Category 3	intercept	ψ_{23}	-1.9772***	-0.0093***	0.0147	0.0301	0.0009	0.0462	-0.0007
			(0.2150)	(0.0025)	(0.1687)	(0.1657)	(0.0007)	(0.1041)	(0.0041)
Category 4	intercept	ψ_{24}	-1.5526***	-0.0104***	0.0517	-0.1309	-0.0001	0.0152	-0.0018
			(0.1796)	(0.0019)	(0.1432)	(0.1498)	(0.0007)	(0.0902)	(0.0036)
Category 5	intercept	ψ_{25}	-1.3386***	-0.0106***	-0.1690	-0.1190	-0.0012	-0.0718	-0.0052
			(0.1454)	(0.0019)	(0.1362)	(0.1262)	(0.0006)	(0.0794)	(0.0032)
Category 6	intercept	ψ_{26}	-4.3391***	-0.0145*	-0.7419	0.5344	-0.0029	0.6925	-0.0088
			(0.4761)	(0.0058)	(0.3510)	(0.4992)	(0.0022)	(0.2761)	(0.0110)
Channel utility parameters									
Retail utility	intercept	ω_2	3.8234***	-0.0021	0.6815**	0.1744	0.0006	0.1343	-0.0140
			(0.2398)	(0.0034)	(0.1933)	(0.2274)	(0.0010)	(0.1351)	(0.0047)
Shopping budget parameters									
Budget	intercept	τ	5.0015***	0.0110***	0.1064	0.0577	0.0008	-0.0067	-0.0008
			(0.1267)	(0.0013)	(0.1065)	(0.1184)	(0.0005)	(0.0721)	(0.0026)
Consideration arrival parameters									
Arrival rate	intercept	δ	-4.2092***	-0.0011	0.0070	-0.2072	0.0001	0.0088	-0.0018
			(0.1192)	(0.0013)	(0.0878)	(0.0877)	(0.0004)	(0.0479)	(0.0021)

* p<0.05, ** p<0.01, *** p<0.001

Table 14: Model hyperparameter estimates

	ψ_{11}	ψ_{12}	ψ_{13}	ψ_{14}	ψ_{15}	ψ_{16}	ψ_{21}	ψ_{22}	ψ_{23}	ψ_{24}	ψ_{25}	ψ_{26}	τ	ω	δ
ψ_{11}	1														
ψ_{12}	-0.013	1													
ψ_{13}	0.152	0.249	1												
ψ_{14}	0.070	0.345	0.339	1											
ψ_{15}	0.079	0.306	0.218	0.328	1										
ψ_{16}	0.104	0.469	0.323	0.493	0.476	1									
ψ_{11}	-0.028	-0.141	-0.164	-0.123	-0.119	-0.137	1								
ψ_{12}	0.002	0.276	0.097	0.177	0.105	0.195	-0.0411	1							
ψ_{13}	0.015	0.123	0.080	0.154	0.053	0.077	0.065	0.103	1						
ψ_{14}	0.018	0.168	0.132	0.248	0.126	0.231	-0.049	0.029	0.114	1					
ψ_{15}	0.026	0.244	0.100	0.217	0.329	0.334	-0.104	0.038	-0.040	-0.010	1				
ψ_{16}	0.130	0.442	0.308	0.439	0.383	0.411	-0.090	0.194	0.116	0.198	0.262	1			
τ	-0.252	-0.240	-0.217	-0.375	-0.326	-0.304	-0.153	-0.202	-0.134	-0.296	-0.433	-0.242	1		
ω	-0.228	-0.104	-0.358	-0.141	-0.174	-0.259	0.323	0.174	0.216	0.149	0.101	0.007	0.002	1	
δ	-0.140	-0.014	-0.154	-0.025	-0.151	-0.137	-0.164	-0.063	-0.161	-0.031	-0.166	0.011	0.331	-0.145	1

Table 15: Correlations among heterogeneous parameters

D Model fit exhibits

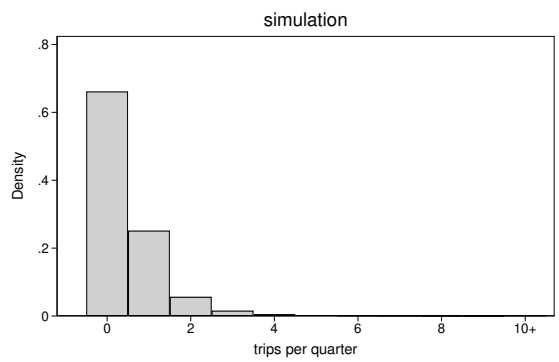


Figure 21: Simulated frequency of purchase occasions per quarter

Moment	Data	Simulation
purchases (#/qtr)	0.51	0.55
online expenditures ($/qtr)	34.32	36.04
retail expenditures ($/qtr)	36.58	42.72
online share of expenditures	54.37%	56.29%
online share of purchases	44.07%	39.58%
online expenditure/purchase ($)	142.45	178.87
retail expenditure/purchase ($)	119.63	138.91

Table 16: Data vs. model-simulated moments

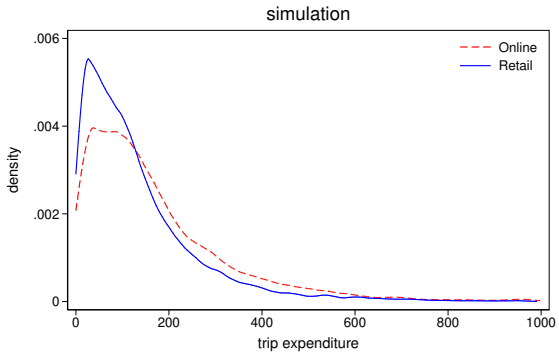


Figure 22: Simulated expenditure distributions by channel

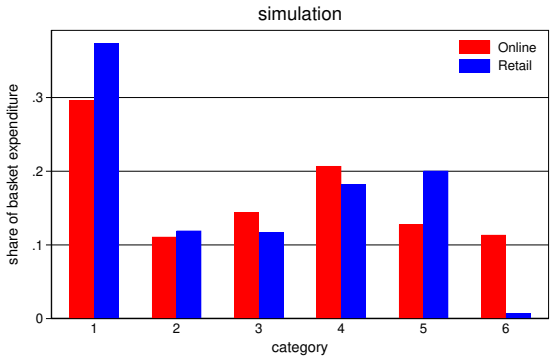


Figure 23: Simulated average category shares by channel

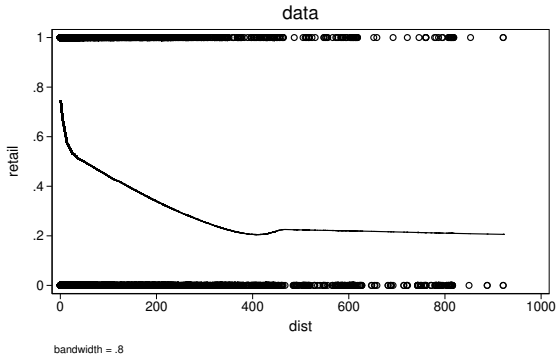


Figure 24: Retail visit probability as a function of store distance, data

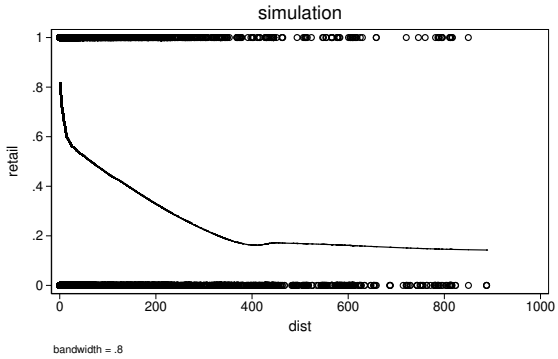


Figure 25: Retail visit probability as a function of store distance, simulation