Structural Analysis of Multi-Channel Demand

Scott Shriver and Bryan Bollinger *

July 15, 2017

Abstract

Not only has the physical retail channel survived the advent of online and mobile retailing, many online retailers are now opening brick-and-mortar stores. These trends can be partially explained by recognizing that different channels serve different roles and deliver distinct consumer experiences. Physical retail outlets not only allow consumers to assess product match through direct product interaction, the presence of a nearby retail outlet may also increase consumer consideration of the brand by creating top-of-mind awareness. Using detailed data on consumer locations and purchase histories during a period of rapid retail expansion, we separately identify the effects of retail store distance on brand consideration (which can increase demand in both channels) and channel switching. Using a direct utility framework that accommodates multi-channel, multi-category demand as well as fixed and variable channel costs, we find that a 10% reduction in distance to the nearest retail store increases total expenditures among existing customers by 1.97%. Revenue gains result from an awareness-induced increase in purchase frequency (1.95%) and partial recapture of consumer transportation cost savings. Retail expenditures rise more rapidly (3.35%) than the number of retail purchases (3.20%), reflecting higher expenditures per retail purchase due to cost savings. As the retail share of lower-budget purchases grows, online expenditures per purchase also rise such that, despite a decline in the number of online purchases (-0.03%), total online revenues also increase 0.07%. For the modal customer, our heterogeneous estimates imply that within five miles of a store, purchase frequency effects dominate channel substitution effects such that online revenues increase as retail distance decreases. In counterfactual experiments, we explore channel-based pricing policies and identify desirable locations for retail entry.

*Shriver: Leeds School of Business, University of Colorado, scott.shriver@colorado.edu. Bollinger: The Fuqua School of Business, Duke University, bryan.bollinger@duke.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, 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 greater 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 awareness of the brand, effectively serving as a form of advertising that can affect demand across multiple channels. 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 In this paper, we develop a unified utility framework to measure these demand-expanding effects of retail entry, as well as the extent to which existing online demand is cannibalized. Our model explains a comprehensive set of demand outcomes including the frequency with which consumers shop, whether they buy from the online (web) or retail channel, and how they allocate expenditures among multiple product categories.

We estimate the model using purchase histories of approximately 10,000 randomly selected customers from a firm that sells directly to consumers through online and retail channels. Customer home locations are observed at the Census block level, yielding highly accurate measures of customer to retail store distances. Critically, we observe significant within-customer variation in retail store distance because the firm doubled its retail footprint over our two-year observation window. We leverage this individual-level variation to identify the effects of retail proximity on purchase frequency and channel expenditure patterns.

Using the observed distribution of customer/store locations, the total (online and retail, per quarter) expenditure elasticity implies a 10% decrease in retail distance increases total expenditures among existing customers by 1.97%. These revenue gains result from an awareness-induced increase in purchase frequency (1.95%) and partial recapture of consumer transportation cost savings. Channel switching patterns imply the frequency of retail purchases increases 3.35% while the frequency of online purchases declines 0.03%. The reduction of retail transportation costs leads to a direct increase in expenditures per retail purchase, as reflected by the fact that retail expenditures rise more rapidly (3.35%) than the number of retail purchases (3.20%). Expenditures per online purchase also rise as lower-budget trips shift to the retail channel, as seen by online expenditures increasing by 0.07% while the number of online purchases declines 0.03%. Our analysis of individual-level estimates reveals that within five miles of a retail store, purchase frequency effects dominate channel switching effects, such that within this radius, the modal consumer’s online revenues increase with proximity to the store.

The rest of the paper is organized as follows: in Section 1.1 we briefly discuss the related empirical literature. In Section 2 we describe the data used for the study and explore variation in key relationships. In Section 3 we develop our structural model. Section 4 describes our estimation method while model estimates are presented in Section 5. We demonstrate the application of our results to problems of retail entry location selection and setting channel pricing policies in Section 6. We summarize our findings and propose further avenues of research in Section 7.

1.1 Related literature

The economic contribution of channels has primarily been explored within the literature on new channel introduction.² Researchers have studied the effect of channel introduction in multiple industries including newspapers

²See Neslin et al. (2006) for a comprehensive survey of the broader literature on multi-channel customer management.
(Deleersnyder et al., 2002; Gentzkow, 2007), music (Biyalogorsky and Naik, 2003) and apparel (Ansari et al., 2008). Most extant studies source data from the period of rapid e-commerce expansion (late 1990s to early 2000s) and consequently investigate the impact of adding an online (website) channel to existing brick and mortar or catalog channels (Ansari et al., 2008; Biyalogorsky and Naik, 2003; Deleersnyder et al., 2002; Geyskens et al., 2002; Van Nierop et al., 2011). Generally, these papers find that the addition of a website does not cannibalize existing 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. To our knowledge, only three other recent studies have analyzed the impact of a firm adding physical stores to its existing online and catalog channels. Avery et al. (2012) use market-level panel data from an apparel/home furnishings retailer to measure the impact of opening a new physical store on net catalog sales, net online sales, sales from new customers and sales from existing customers. They use a differences-in-differences methodology, where control markets are identified using a propensity scoring algorithm, finding that both catalog and online sales are cannibalized in the short run but tend to recover over time.

Pauwels and Neslin (2015) use vector autoregression to analyze similar data (also apparel categories) in an aggregate time series format. The VAR approach permits Pauwels and Neslin (2015) 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 results are generally consistent with Avery et al. (2012) in that the authors find physical store introduction: i) cannibalizes catalog sales but not online sales and ii) increases the rate of new customer acquisition.

The third paper is Wang and Goldfarb (2017), which uses some of the same data we analyze. Wang and Goldfarb (2017) also use a differences-in-differences methodology to investigate the net impact of retail entry on online and retail sales in markets defined by Census tracts. Wang and Goldfarb (2017) find that retail entry drives new customer acquisition in both the online and retail channels. The authors further find that in areas with a weak brand presence (defined as the Census tract having zero sales in the three months prior to entry), the customer acquisition effect dominates as there are no pre-existing online sales to cannibalize. In other areas, online channel sales are partially cannibalized in response to store entry.

Our empirical findings are broadly consistent with the previous literature in that all studies find evidence that retail entry leads to demand expansion through new customer acquisition and most (all but Pauwels and Neslin (2015)) find evidence of online channel cannibalization. Our analysis is differentiated by being the only one conducted at the level of individual customer purchase transactions, which confers several benefits. First, we avoid potential aggregation bias associated with using market-level revenue data. This is especially a concern in estimating and interpreting retail distance demand effects, since total transportation costs will depend on the number of retail trips made during the period in question. Second, we can exploit within-customer demand responses to nearby store entry to identify effects of interest. That is, we use covariation in demand and retail store distance to identify effects, after implementing rich controls for unobserved consumer heterogeneity. Relatedly, we conjecture that our operationalization of retail channel effects using individual customer-to-store distances provides a richer source of identifying variation than counts of stores within a geographic market (as used in the other studies). Finally, our structural model is the only framework that controls for channel-specific product prices and offers rich decision support capabilities to forecast demand under counterfactual price and retail distance regimes.

The setup of our model also links it to the 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. In particular, the ability to assess product fit information has been shown to be an important factor in channel format choices in several recent marketing studies (e.g., Bell et al., 2013; Soysal and Zentner, 2014; Dzyabura et al., 2015). Our model captures these information differences in product and channel specific shocks to consumer utility. 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 trip budget and channel transaction costs. Whereas the standard practice in the store choice literature has been to treat expenditure levels as exogenously determined,\(^4\) we endogenize the expenditure decision as well as the channel choice decision.

2 Data

2.1 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).\(^5\) 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.\(^6\)

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 2.4, 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.

\(^4\)For example, Bell and Lattin (1998) use a pre-estimation calibration to obtain the probability that consumers are either “large basket” or “small basket” shoppers, and Bell, Ho, and Tang (1998) assume budgets arise implicitly from an unobserved shopping list.

\(^5\)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.

\(^6\)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.
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).\(^7\)

### 2.2 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 use a discrete time formulation and organize the data as a collection of customer purchase occasions within a time period (quarter). That is, each period a customer is observed (the customer’s first purchase quarter and subsequent quarters), may be associated with zero or more purchases where each purchase is specific to one channel. This approach allows us to associate continuously-varying price and store distance information with summary discrete-time measures.

Third, we abstract from SKU-level choices and characterize purchase outcomes in terms of: a) the total expenditure level (excluding taxes and shipping fees), b) the purchase channel indicator, and c) expenditure shares 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) but 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 to demand, we omit returned items from our computation of expenditure levels and category shares.

### 2.3 Summary

We first discuss variation in the data at the customer/quarter panel level, and then turn to discussion of variation at the level of individual purchase occasions. Table 1 summarizes variables with panel-level variation (number of purchases, quarterly expenditures by channel, retail store distance) and with cross-sectional variation (demographics, taxes). The panel summarized in Table 1 is unbalanced across the 8 quarters of study and 10,239 individuals, reflecting the observed acquisition of new customers.

The table indicates 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 and retail 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. The within-customer variation induced by observed retail entry events is substantial, with a standard deviation of 28.4 miles.

---

\(^7\)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://medc.missouri.edu/websas/geocorr12.html.
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.

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

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

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:

$$p_{kct} = \exp \left( \frac{\sum_{j \in J_t, k} w_j \log(p_{jct})}{\sum_{j \in J_t, k} w_j} \right)$$

where: $$w_j = \frac{\sum_{c} \sum_{t} e_{jct}}{\sum_{c} \sum_{t} e_{jct}}$$

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](image)

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.

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

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 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.
2.4 Descriptive analyses

In this section, we explore how distance to the retail outlet affects demand along four dimensions: purchase frequency, the total expenditure per purchase, the channel chosen for purchase, and new customer acquisition. Our findings here guide the formulation of our structural model in Section 3 and related extensions in Section 6.1.

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

2.4.1 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 (\( \delta_i, \mu_t \)) to control for unobservables – the effect of retail store distance (\( \alpha \)) 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:

\[
L_{it} \sim \text{Poisson}(\rho_{it}), \quad \log(\rho_{it}) = \alpha \log(d_{it}) + \delta_i + \mu_t
\]

The estimate of \( \alpha \) 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 \( p \) 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, presumably due to increased consideration of the brand (both channels) and reduced transportation costs (retail channel).

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 identifies the distance effects using within-household variation by displaying the predicted values from regressions in the form of equation (2), conditioned upon retail (panel a) and online (panel b) purchases only, and replacing the \( \log(d_{it}) \) term 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. However, compared to retail purchases, increases in online purchases only occur in quite close proximity to retail locations.
2.4.2 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. 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 ($\delta_i$, $\mu_t$):

$$\log(e_{itl}) = \alpha \log(d_{it}) + \delta_i + \mu_t + \epsilon_{itl}$$  (3)

The estimate of $\alpha$ 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.

In the raw data, plotted in the left side graphs in Figure 7, there is no clear relationship between expenditure per purchase and distance. As before, we estimate regressions using semi-parametric distance variants of equation (3) and plot these in the right hand graphs in Figure 7. These plots also suggest no obvious relationship between expenditure per purchase and distance. We conclude there is little evidence to support modeling a direct dependence of shopping budget (hence expenditure) levels on retail distance.\(^8\)

2.4.3 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

\(^8\)We expect (and our model allows for) expenditure levels to depend indirectly on retail distance through transportation costs, which reduce the budget available for retail purchases. These transportation costs potentially affect both the purchase expenditure level and which channel is chosen, where the latter effect is shown to be significant in Section 2.4.3 below.
**Figure 7: Distance effect on customer expenditure per purchase**

(a) Retail

(b) Online

Note: Error bars designate 95% confidence intervals. In the right side graphs, the upper value of the bin is displayed on the axis.

**Figure 8: Distance effect on retail channel choice frequency**

Note: Error bars designate 95% confidence intervals. In the right side graphs, the upper value of the bin is displayed on the axis.

The estimated $\hat{\alpha}$ is $-0.531$ 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 due to store entry increase use of the retail channel. As before, Figure 8 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 a channel choice dependence on retail distance.
2.4.4 New customer acquisition

Finally, we consider the impact of retail distance on new customer acquisition. Although our structural framework is principally focused on existing customer demand, in Section 6.1 we demonstrate how it may be extended to incorporate new customer acquisition.

To assess how retail store distance influences new customer acquisition, we construct an auxiliary dataset organized as a balanced panel of 8 quarterly observations on 216,291 Census block groups, summarized in Table 3. Variables include the distance from the Census block group centroid to the nearest retail store ($d$), the number of new ($N$) and existing ($E$) customers, and the market population ($pop$). In total, 7641 customers are acquired over the course of the observation window (75% of individuals in our sample), leading to substantial panel-level variation in the variable $N$.

<table>
<thead>
<tr>
<th>observations</th>
<th>mean</th>
<th>std dev</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>retail distance (mi) ($d$)</td>
<td>1,730,328</td>
<td>147.758</td>
<td>163.653</td>
<td>0.01</td>
</tr>
<tr>
<td>new customers ($N$)</td>
<td>1,730,328</td>
<td>0.004</td>
<td>0.067</td>
<td>0</td>
</tr>
<tr>
<td>existing customers ($E$)</td>
<td>1,730,328</td>
<td>0.020</td>
<td>0.150</td>
<td>0</td>
</tr>
<tr>
<td>population ($pop$)</td>
<td>1,730,328</td>
<td>1407.82</td>
<td>811.03</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3: New customer data (Census block group/quarter panel)

To assess how the presence of a retail outlet may influence new customer acquisition, we model the arrival of new customers ($N_{mt}$) as a Poisson process:

$$N_{mt} \sim \text{Poisson}(\rho_{mt}), \quad \log(\rho_{mt}) = \alpha \log(d_{mt}) + \delta_m + \mu_t + \log(pop_m - E_{mt})$$  \hspace{1cm} (5)

As in Section 2.4, the specification includes bi-level (market, quarter) fixed effects ($\delta_m, \mu_t$) to control for unobservables. The level of the arrival rate is normalized in proportion to the number of consumers in a market who are not yet customers of the brand ($pop_m - E_{mt}$). The estimate of $\alpha$ from this regression is $\hat{\alpha} = -0.356$ with a standard error of 0.022 (highly significant). Thus, a 10% decrease in retail store distance corresponds to a 3.6% increase in the mean number of new customers per quarter. The relatively large magnitude of this effect indicates that new customer acquisition is a critical economic benefit of operating a retail channel. The finding supports the notion that retail stores are effective in boosting initial brand awareness among prospective customers.

In summary, the descriptive analyses suggest two key effects of retail proximity on existing customer demand: an effect on the overall shopping frequency and an effect driving channel choice. We therefore include direct dependencies on store distance for these factors when specifying in our structural model in Section 3 below.

3 Model

Our modeling objective is to use economic primitives to explain and predict a comprehensive set of demand outcomes: how frequently consumers purchase, their choice of purchase channel, how much they spend, and their allocation of expenditures 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 3.6.

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 rate parameter $\mu$. 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. As described
in more detail in Section 3.5, 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 one of the brand’s channels.

Conditional upon a brand consideration event, the consumer allocates the total shopping budget \( b \) among channels, product categories and the outside option in a two-stage budgeting decision.\(^{10}\) In the first stage, the consumer chooses the expenditure level in dollars for each channel \( e_c \) such that, at most, one channel has positive expenditure for a given consideration event. This choice includes the possibility of zero expenditures in all channels (a no-purchase condition). Channel expenditure levels are constrained such that the total shopping budget is equal to the sum of: (i) channel expenditures (plus applicable sales taxes), (ii) the outside good expenditure, and (iii) channel fixed costs. Channel fixed costs include shipping costs (online) and transportation costs (retail). In the second stage, the consumer allocates the channel expenditure among the \( K \) categories by setting expenditure shares \( s_k \), where the \( s_k \) sum to one.

In the following subsections, we first formulate the consumer’s decision problem (Section 3.1) and then discuss the empirical specifications we use during model estimation (Section 3.2). Next, we solve the model and derive likelihood expressions for each stage in Sections 3.3, 3.4 and 3.5. Finally, we introduce our persistent heterogeneity specification in Section 3.6.

### 3.1 Consumer’s problem: stages 1 and 2

We collect a consumer’s channel expenditure choices in a \( C \times 1 \) vector \( \tilde{e} \) and her category share choices in a \( K \times 1 \) vector \( \tilde{s} \). Conditional upon a brand consideration event, the consumer’s problem is to select the channel expenditure levels and category shares that maximize her utility, subject to a budget constraint and the restriction that a purchase cannot involve multiple firm channels. If the consumer were to make these decisions jointly, her

---

\(^{10}\) We warrant 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. In this sense, our model serves as a baseline against which other approaches may be compared.
problem can be formulated as follows:

$$\max_{\tilde{\bar{e}}, \tilde{s}} U(\tilde{\bar{e}}, \tilde{s}) = \sum_{c=1}^{C} \sum_{k=1}^{K} \Psi_{ck} Y_k \left( \frac{s_k e_c}{p_{ck} Y_k} + 1 \right) + \sum_{c=1}^{C} \Omega_c I(e_c > 0) + \log(z)$$  \hspace{1cm} (6a)

subject to:  
\[ e_c \geq 0, \sum_{c=1}^{C} I(e_c > 0) \in \{0,1\}, z > 0 \]  \hspace{1cm} (6b)

\[ s_k \geq 0, \sum_{k=1}^{K} s_k = \sum_{c=1}^{C} I(e_c > 0) \]  \hspace{1cm} (6c)

\[ \sum_{c=1}^{C} e_c (1 + r_c) + \sum_{c=1}^{C} f_c I(e_c > 0) + z = b \]  \hspace{1cm} (6d)

The first term in the specification in (6a) above captures the utility consumers get from the purchase of products from the focal brand. Note that \( s_k e_c / p_{ck} \equiv q_{ck} \) is a quantity index for purchases in category \( k \) and channel \( c \), so that consumers have decreasing marginal utility in the quantity of each category (through the \( \log(\cdot) \) expression). The addition of one to each term on the right-hand side of (6a) inside the \( \log(\cdot) \) expression enforces a lower bound of zero for each of the additively separable category/channel sub-utilities. \( \Psi_{ck} \) is the “baseline” marginal utilities (whose stochastic empirical specifications are shown in Table 4), because \( \Psi_{ck} \) is the marginal utility obtained in the limit of consuming zero quantity (\( \lim_{q_{ck} \rightarrow 0} \frac{\partial U}{\partial q_{ck}} \)).

The second term in (6a) captures utility that stems from the shopping activity itself. That is, \( \Omega_c \) is the utility obtained from the act of shopping in channel \( c \). We include these terms to reflect the fact that some consumers have strong preferences over channel formats, independent of how much they spend or what products they choose to buy. The indicator function \( I(e_c > 0) \) reflects that shopping utility is only incurred for a chosen channel. The final term in (6a) captures the consumer’s utility for the outside good, where \( z \) is the expenditure level for the outside good. 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.

Equation (6b) contains constraints on the channel expenditure levels. The first expression implies expenditures must be non-negative, while the second enforces that at most one channel may be chosen for each consideration event. Equation (6c) provides constraints on category expenditure shares: these too must be non-negative and sum to one if there is a positive expenditure in some category (and must be identically zero otherwise). The final relation (6d) is the overall shopping budget for the consideration event: the budget \( b \) equals channel expenditures (marked up by the channel sales tax rate, \( r_c \)) plus the outside good expenditure \( z \) and channel fixed costs \( (f_c) \). As with the channel shopping utility, channel fixed costs are only incurred when channel expenditures are positive.

An empirical specification of the consumer’s decision model imposes functional form and stochastic assumptions on \( \{\Psi_{ck}, \Omega_c, f_c, b\} \). Under standard assumptions for these terms, computing the optimal channel expenditure levels and category shares in the joint decision model is prohibitively expensive. To simplify computation and to reflect a more realistic information revelation process, we assume that consumers sequentially make channel expenditure and category share allocation decisions. Shocks to channel utilities (contained in \( \Omega_c \)) and the trip budget (contained in \( b \)) are realized at the time of the first stage channel expenditure decision, while shocks to product category utilities (contained in \( \Psi_{ck} \)) are realized at the time of the second stage category share decision. In the first stage, we assume consumers have rational expectations with respect to product prices in both channels and know the distribution of \( 2^{nd} \) stage unobservables. Under these assumptions, expectations are taken with respect to \( \Psi_c \). Let \( v_c(e_c) = E_{\Psi_c} \left[ \sum_{k=1}^{K} \Psi_{ck} Y_k \log \left( s_k \frac{\Psi_{ec}}{p_{ck} Y_k} + 1 \right) \right] \) be the expected utility from expenditure \( e_c \) in outside good (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.

\[^{11}\text{Imposing diminishing marginal utility for the outside good } z \text{ (though the } \log(\cdot) \text{ expression) is essential in our context. Assuming linear utility in the outside good } z \text{ would result in optimal demand conditions that do not involve the shopping budget or channel fixed costs, which are key objects of inference.}\]
channel \( c \), assuming categories are optimally chosen (denoted by the star superscript on \( s_k \)). Then, the two stage budgeting decision may be written as:

**Stage 1: Channel expenditure decision**

\[
\max_{\bar{\xi}} V(\bar{\xi}) = \sum_{c=1}^{C} v_c(e_c) + \sum_{c=1}^{C} \Omega_c I(e_c > 0) + \log(z) \tag{7a}
\]

subject to: \( e_c \geq 0, \sum_{c=1}^{C} I(e_c > 0) \in \{0, 1\}, z > 0 \tag{7b} \)

\[
\sum_{c=1}^{C} e_c(1 + r_c) + \sum_{c=1}^{C} f_c I(e_c > 0) + z = b \tag{7c}
\]

where: \( v_c(e_c) = E_{\Psi_c} \left[ \sum_{k=1}^{K} \Psi_{ck} \gamma_k \log \left( \frac{s_k^{\psi}(\Psi_c)e_c}{p_{ck}y_k} + 1 \right) \right] \tag{7d} \)

**Stage 2: Category share decision**

\[
\max_{\bar{s}} u(\bar{s} | e_c^* > 0) = \sum_{k=1}^{K} \Psi_{ck} \gamma_k \log \left( \frac{s_k e_c^*}{p_{ck}y_k} + 1 \right) \tag{8a}
\]

subject to: \( s_k \geq 0, \sum_{k=1}^{K} s_k = 1 \tag{8b} \)

We discuss computational issues related to solving this two-stage budgeting model in Section 4.2 and Appendix A.

### 3.2 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 4 below.

<table>
<thead>
<tr>
<th>Functional form</th>
<th>Distributional assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category utilities</td>
<td>( \Psi_{ck} = \exp(\psi_k + \xi_c + \eta_{ck}) )</td>
</tr>
<tr>
<td>Channel utilities</td>
<td>( \Omega_c = \alpha_k + \xi_c )</td>
</tr>
<tr>
<td>Channel fixed costs*</td>
<td>( f_1 = 7.95, f_2 = \alpha d^\xi )</td>
</tr>
<tr>
<td>Trip budget</td>
<td>( b = \exp(\tau + \beta^T X^T + \nu) )</td>
</tr>
<tr>
<td>Consideration arrival*</td>
<td>( \mu = \exp(\delta + \beta^T X^T) \left( 1 + (1 + d^\xi) \right) )</td>
</tr>
</tbody>
</table>

* \( d \) = retail store distance

Table 4: Empirical specifications for \( \{\Psi_{ck}, \Omega_c, f_c, b, \mu\} \)

Category marginal utilities \( \Psi_{ck} \) are constrained to be positive through the \( \exp(\cdot) \) operator, and are comprised of a category intercept (\( \psi \)) and two stochastic terms (\( \eta \) and \( \epsilon \)). The \( \eta \) terms are normally distributed channel/category specific shocks that capture product fit and assortment information. By assumption, the extreme value shocks (\( \epsilon \)) are iid across categories and channels. This compound shock specification is used to allow for flexible shock variance structures while providing computational convenience when evaluating the likelihood function.\(^{12}\) The interpretation of these structural shocks is that they are the information realized upon visiting

\(^{12}\)In deriving the likelihood function, integration over the iid extreme value shocks (\( \epsilon \)) may be done analytically. Monte Carlo integration is performed over the heteroskedastic normal shocks (\( \eta \)), using draws that serve the dual purpose of computing the expected product utility function \( v_c(e_c) \).
the channel. We expect that both cross-channel product assortment variation (e.g., unobserved stock-outs) and the desirability of interacting with the product in the retail channel contribute to the shock variances. Given the exponential form in Table 4, a higher shock variance yields a higher expected utility, such that larger shock variances are expected in categories with more complete assortments (fewer stock-outs).

The channel utility $\Omega_c$ is comprised of an extreme value shock ($\xi_c$) and an intercept ($\omega_c$) that captures channel format preferences. 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 flexible function of the consumer’s distance ($d$) to the nearest retail store: $f_2 = \alpha d^k$.

We assume a log-normal distribution for the trip budget ($b$). In addition to an intercept ($\tau$), the trip budget specification incorporates additional controls through the shifter matrix $X^\tau$. Our specification of $X^\tau$ includes 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 $\mu$ is specified as a flexible function of retail store distance, an intercept ($\delta$) and additional controls ($X^\delta$). The term incorporating retail store distance $(1 + (1 + d^\delta))$, which captures the effect of retail store proximity on a consumer’s propensity to consider the brand, is constructed such that the arrival rate is bounded as $d \to 0$ and $d \to \infty$ (provided $\xi < 0$, as is expected). In $X^\delta$, we include quarter of year dummies to account for seasonal patterns of shopping frequency.

### 3.3 Second stage optimality conditions and likelihood

**Optimality conditions** For the second-stage sub-problem (allocation of expenditure across categories), both the objective (utility) function and the constraint equation are continuously differentiable in the control variables ($\delta$), so the Kuhn-Tucker theorem may be applied to solve the model. We first form the Lagrangian for the problem: $\mathcal{L} = u(s|\epsilon^c) - \lambda \left( \sum_{k=1}^K s_k - 1 \right)$. Optimal shares will then satisfy the Kuhn-Tucker (KT) conditions:

$$\frac{\partial \mathcal{L}}{\partial s_k} \leq 0, s_k \geq 0, \frac{\partial \mathcal{L}}{\partial s_k} s_k = 0 \text{ for all } k = 1, \ldots, K.$$ Without loss of generality, assume that the first good is chosen. In the Technical Appendix, we show that the KT conditions for the consumer’s problem reduce to:

$$g_{ck} + \epsilon_{ck} = \begin{cases} g_{c1} + \epsilon_{c1} & \text{if } s_k^* > 0 \\ g_{c1} + \epsilon_{c1} & \text{if } s_k^* = 0 \end{cases} \quad (9a)$$

where:

$$g_{ck} = \psi_k + \eta_{ck} - \log \left( \frac{s_k^* \epsilon_c}{p_{ck} \gamma_k} + 1 \right) - \log (p_{ck}) \quad (9b)$$

**Likelihood** The Kuhn-Tucker conditions are used to form the likelihood for the observed share vector. Without loss of generality, let the first $M$ ($M \geq 1$) categories be chosen. Conditional upon $\epsilon_{c1}$ and $\eta_{ck}$, chosen categories generate equality constraints that provide a 1:1 mapping from observed (optimal) shares $s_k$ to the shocks $\epsilon_{ck}$. Through a change of variables calculus (Jacobian term), this mapping produces a probability density for non-zero shares. Non-chosen categories generate inequalities that provide an upper bound on the $\epsilon_{ck}$. Integrating the joint probability density over the allowed regions for $\epsilon_{ck}$ produces a probability for zero share categories. The unconditional likelihood is then given by integrating over $\epsilon_{c1}$ and $\eta_{ck}$. We partition the iid channel/category shocks $\tilde{\epsilon}$ in two parts, with the $Mx1$ vector $\tilde{\epsilon}_M$ corresponding to chosen alternatives and the $(K-M)x1$ vector $\tilde{\epsilon}_{\tilde{M}}$ corresponding to non-chosen alternatives. With this convention, the probability of observing expenditure share pattern $\tilde{s}$ may be written:

$$\ell(\tilde{s} = \tilde{s}^*|\epsilon^c) = \int_{\mathcal{E}M} \int_{\eta} \int_{\epsilon_1}^{\infty} \int_{\epsilon_1}^{\infty} \cdots \int_{\epsilon_1}^{\infty} f(\tilde{\epsilon}) f(\tilde{\eta}) d\tilde{\epsilon}_M d\tilde{\eta} \quad (10)$$

where $f(\cdot)$ represents a probability density function (extreme value for $\epsilon$, normal for $\eta$) and $g$ is defined in equation (9b).
In the Technical Appendix, we show the resulting likelihood reduces to:

\[
\ell(\tilde{s} = \tilde{s}^* | e^*_c) = \frac{(M - 1)!}{\sigma_{ec}^{M-1}} \left( \prod_{j=1}^{M} \frac{e^*_c}{\sum_{j=1}^{M} s_j^* e^*_c + \gamma_j p_{cj}} \right) \left( \sum_{j=1}^{M} \frac{g_{cj}}{\sigma_{ec}} \right) \left( \sum_{j=1}^{K} \frac{g_{cj}}{\sigma_{ec}} \right)^{-M} \phi(\tilde{\eta})d\tilde{\eta}
\]

where \( \phi(\cdot) \) is the normal pdf. The integration over \( \eta \) has no closed form solution – we employ simulation methods to compute this integral in our estimation procedure.

### 3.4 First stage optimality conditions and likelihood

**Optimality conditions** In the first-stage sub-problem (setting channel expenditures), the Kuhn-Tucker theorem does not formally apply because neither the objective function nor the budget constraint is continuously differentiable in the control variables \( \tilde{e} \).

Fortunately, the structure of the problem presents a viable solution procedure. To fix ideas, consider our empirical setting where there are two channels. Note that because at most one channel will receive positive expenditure, the optimal expenditure vector \( \tilde{e}^* \) must be restricted as follows:

\[
\tilde{e}^* \in \left\{ \left( \begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ e_i^* \\ 0 \\ e_j^* \\ \vdots \\ 0 \\ e_2^* \end{array} \right) : \{ e_0^*, e_1^*, e_2^* \} \right\} = \{ e_0^*, e_1^*, e_2^* \},
\]

where \( e_i^* \) and \( e_j^* \) are scalar values that represent the optimal expenditure levels for channels 1 and 2, assuming they were the chosen channel. More generally, in a slight abuse of notation, let:

\[
\tilde{e}_c^* \equiv Cx1 \text{ vector with } j^{th} \text{ element } e_j I(c = j) , \quad \tilde{e}_0^* \equiv 0_{C \times 1}
\]

With this convention, an observation of \( \tilde{e}_c^* \) corresponds to the following two optimality conditions:

1. \( e_i^* \) solves \( \frac{\partial V(\tilde{e}_c)}{\partial e_i} = 0 \) : optimal trade-off between expenditures in chosen channel and the outside good.

2. \( V(\tilde{e}_c) = \max_{j \neq c} V(\tilde{e}_j) \) : chosen channel utility exceeds (expenditure optimal) counterfactual channel, outside good only utilities.

Note that for the chosen channel, the optimal expenditure level (\( e_i^* \)) is observed. However, assessing the second optimality condition requires us to compute (as described below) the optimal expenditure level (\( e_j^* \)) for each unchosen channel.

**Optimality condition 1** Given that channel \( c \) is chosen, terms from the alternative channels drop out of the utility and budget constraint equations. By solving the budget constraint for \( z \) and substituting into the utility function \( V \), the first optimality condition may be evaluated as:

\[
\frac{\partial V(\tilde{e}_c)}{\partial e_c} = \frac{\partial}{\partial e_c} \left( v_c(e_c) + \Omega_c + log \left( b - e_c(1 + r_c) - f_c \right) \right) = 0.
\]

Solving the resulting first order condition for the budget \( b \) gives:

\[
b = f_c + (1 + r_c) e_i^* + \frac{1 + r_c}{v'_c(e_i^*)}
\]

Comparing equations (13) and (7c) reveals that the first order condition pins down the outside good expenditure, \( z = \frac{1 + r_c}{v'_c(e_i^*)} \), and hence the trip budget \( b \), as a function of the model parameters and observable data. Using our empirical specification for \( b \) from Table (4), taking logs and rearranging gives an optimality condition for the shock to the trip budget, \( v \):

\[
v = h(e_i^*) = log \left( f_c + (1 + r_c) \left( e_i^* + \frac{1}{v'_c(e_i^*)} \right) \right) - \tau - \beta^b X^b
\]

\[\text{Footnote: Indicator functions in the channel utility and budget constraint equations produce discontinuities in the derivatives when moving from zero to positive expenditure.}\]
Optimality condition 2  First, for notational convenience, we let $\bar{V}_c$ represent the consumer’s total expected utility at her optimal expenditure level in channel $c$:

$$
\bar{V}_c \equiv E\left[V(c^*_{c})\right] = v_c(e^*_c) + \omega_k + \log(z) = v_c(e^*_c) + \omega_k + \log\left(\frac{1+r_c}{v'_c(e^*_c)}\right)
$$

(15)

where we use $z = \frac{1+r_c}{v'_c(e^*_c)}$ from optimality condition 1 for the second relation. For $V(c^*_{c}) = \max_{j=0,...,C}\left[V(c^*_{j})\right]$ to hold, the utility of consuming in channel $c$ must exceed the utility of consuming the optimal amount in any other channel and that of allocating the entire budget to the outside good:

$$
V(c^*_{c}) = \bar{V}_c^* + \xi_c \geq V(c^*_{0}) = \log(b)
$$

(16a)

$$
V(c^*_{j}) = \bar{V}_c^* + \xi_c \geq V(c^*_{j}) = \bar{V}_j^* + \xi_c , \; j \neq c
$$

(16b)

Both $\bar{V}_c^*$ and $b$ in condition (16a) are expressible in terms of model parameters and observable data. However, the optimal expenditure levels in the unchosen channels, $e^*_j$, must be computed in order to generate $\bar{V}_j^* = \bar{V}_j^*(e^*_j)$.

To do this, we find the $e^*_j$ that solves the budget constraint for channel $j$: $f_j + (1+r_j)\left(e_j + \frac{1}{v'_j(e_j)}\right) = b$, where the value of $b$ is substituted from equation (13) above. $e_j^*$ may be found numerically using bisection algorithms or other root finding techniques. With the $e^*_j$ values in hand, the $\bar{V}_j^*$ terms may be computed using equation (15).

Likelihood  As with category shares, the joint likelihood of the expenditure vector is formed by integrating the joint density of the budget and channel shocks over the region consistent with the optimality conditions:

$$
\ell(\tilde{c} = \bar{c}^*_{c}) = |J_{\bar{V} \rightarrow e^*_c}| f_V(h(e^*_c)) \int_{-\infty}^{\infty} \left(V(\bar{c}^*_{c}) = \max_{j=0,...,C}\left[V(\bar{c}^*_{j})\right]\right) f(\tilde{\xi}) d\tilde{\xi}
$$

(17a)

$$
= |J_{\bar{V} \rightarrow e^*_c}| \phi\left(\frac{h(e^*_c)}{\sigma_V}\right) \int_{\xi_v = \log(b) - \bar{V}^*_c}^{\infty} \prod_{j \neq c} \int_{\xi_j = -\infty}^{\bar{V}^*_j - \bar{V}^*_c + \xi_c} \lambda\left(\frac{\bar{V}^*_j}{\sigma_{\xi_j}}\right) d\xi_j \lambda\left(\frac{\bar{V}^*_c}{\sigma_{\xi_c}}\right) d\xi_c
$$

(17b)

where $\lambda(\cdot)$ is the extreme value pdf and equation (17b) makes use of our empirical specification distributional assumptions on $v$ and $\xi$.. In the Technical Appendix, we show the integration in equation (17b) may be performed analytically, resulting in the following expression for the likelihood of the expenditure vector:

$$
\ell(\tilde{c} = \bar{c}^*_{c}) = |J_{\bar{V} \rightarrow e^*_c}| \phi\left(\frac{h(e^*_c)}{\sigma_V}\right) \left(\frac{\exp\left(\frac{\bar{V}^*_c}{\sigma_{\xi_c}}\right)}{\sum_{j=1}^{C} \exp\left(\frac{\bar{V}^*_j}{\sigma_{\xi_j}}\right)}\right) \left(1 - \frac{C}{\prod_{j=1}^{C} \lambda\left(\frac{\log(b) - \bar{V}^*_j}{\sigma_{\xi_j}}\right)\right)
$$

(18)

where $h(\cdot)$, $\bar{V}^*_c(\cdot)$ and $b(\cdot)$ are defined above in (14), (15) and (13) and $\Lambda(\cdot)$ is the extreme value cdf. The final term in equation (18) corresponds to the probability of purchase (not allocating all expenditure to the outside good), while the middle term corresponds to the probability of the channel being chosen conditional on purchase. The first term reflects the expenditure level probability density conditional upon the chosen channel.

---

14 The solution for $e^*_j$, if it exists, is unique. To see this, first rewrite the budget constraint as $e_j = \frac{b-f_j}{1+r_j} - \frac{1+r_j}{v'_j(e_j)}$ and note that $v'_j(e_j)$ is both positive and decreasing in $e_j$. Therefore, the left hand side of this equation is monotonically increasing in $e_j$ while the right hand side is monotonically decreasing: if such functions intersect, they do so at a single point, implying any solution $e^*_j$ must be unique. As $e_j$ is bounded on $[0, \frac{b-f_j}{1+r_j})$, sufficient conditions for existence are that $0 \leq \frac{b-f_j}{1+r_j} - \frac{1+r_j}{v'_j(e_j)}$ and $\frac{b-f_j}{1+r_j} > \frac{b-f_j}{1+r_j} - \frac{1+r_j}{v'_j(e_j)}$. The latter condition is trivially satisfied while the first holds when the residual budget after fixed costs and taxes $\left(\frac{b-f_j}{1+r_j}\right)$ exceeds the outside good expenditure at zero channel expenditure $\left(\frac{1+r_j}{v'_j(0)}\right)$. The condition fails, for example, when counterfactual channel fixed costs exceed the budget. In such cases, we set $\bar{V}^*_j = -\infty$ when evaluating the model likelihood, to reflect the fact the channel is not a viable choice at this budget level.

15The Jacobian term is: $J_{\bar{V} \rightarrow e^*_c} = \frac{\partial \bar{V}^*_c}{\partial e^*_c} = \frac{\partial h(e^*_c)}{\partial e^*_c} = (1+r_c)\left(1 - \frac{\left(v'_j(e^*_j)\right)}{\left(v'_j(e^*_c)\right)}\right)\left(f_c + (1+r_c)\left(e^*_c + \frac{1}{v'_j(e^*_c)}\right)\right)^{-1}$

16
3.5 Consideration arrival/purchase incidence likelihood

Let $L$ be the number of observed purchase occasions in the period. We assume that purchases originate from a Poisson arrival process with rate parameter $\mu$, which reflects the number of times per quarter a consumer considers making a purchase with the focal brand. Since our expenditure model admits the possibility of zero expenditures in any channel, not every “consideration event” must result in a purchase. Accordingly, we derive the likelihood of observing $L$ purchase occasions in two steps, where we first quantify the probability that a consideration event results in no purchase. We then integrate over the set of no purchase events that can result in the observed number of purchases.

3.5.1 No purchase probability

Following the logic of optimality condition 2 above and given the independence of the channel shocks, the probability of no purchase given consideration (at a fixed budget level $b$) is:

$$Pr(V^*_0 > V^*_c, \forall c | b) = \prod_{c=1}^{C} \Lambda \left( \frac{\log (b) - \bar{V}^*_c(b)}{\sigma_s} \right)$$

To compute this probability for a given budget, one must solve for the optimal expenditure level in each channel to evaluate $\bar{V}^*_c$ (as in optimality condition 2 above). Then, the unconditional probability of no purchase is given by integrating over the budget distribution:

$$\ell (\bar{e} = \bar{e}_0) = \int \prod_{c=1}^{C} \Lambda \left( \frac{\log (v) - \bar{V}^*_c(v)}{\sigma_s} \right) \phi(v) dv$$

This integral may be computed by Gauss-Hermite quadrature.

3.5.2 Purchase incidence likelihood

To form the likelihood of $L$, we integrate over the distribution of unobserved consideration events that can rationalize an observation of $L$ purchase occasions, as follows:

$$\ell (L) = \sum_{k=0}^{\infty} Poisson (L+k, \mu) \left( \binom{L+k}{k} \ell (\bar{e} = \bar{e}_0)^k \right)$$

$$= \sum_{k=0}^{\infty} e^{-\mu} \frac{\mu^{L+k}}{(L+k)!} \frac{(L+k)!}{L!k!} \left[ \ell (\bar{e} = \bar{e}_0)^k \right] = \frac{e^{-\mu} \mu^{L}}{L!} \cdot e^{\mu \ell (\bar{e} = \bar{e}_0)}$$

The above expression reflects, for example, that 3 purchase events could result from 3 consideration events (with zero no-purchases), 4 consideration events (with 1 no-purchase), 5 consideration events (with 2 no-purchases), and so forth. The presence of the no-purchase probability in equation (21) provides a linkage between the purchase incidence rate and the structural parameters governing channel and product utility.

3.6 Heterogeneity specification

We allow for heterogeneous product and channel preferences by incorporating individual-specific heterogeneity in the category ($\psi$) and channel ($\omega$) utility parameters. We further allow for heterogeneous shopping budgets ($\tau$), retail fixed costs ($\alpha$) and brand consideration rates ($\delta$) via individual-specific parameters. We collect the heterogeneous parameters in the vector $\theta^*_i \equiv \{ \psi_i, \omega_i, \tau_i, \alpha_i, \delta_i \}$. Similarly, we collect the homogeneous parameters in
the vector $\mathbf{\theta}^2 \equiv \{ \gamma, \sigma, \kappa, \beta^*, \beta^\delta \}$. We assume a hierarchical model such that:

$$
\begin{align*}
\theta^1_i & \sim N(Z_iY, \Sigma) \\
\theta^2 & \sim N \left( \theta^2, A_{\theta^2}^{-1} \right) \\
\text{vec}(\Upsilon) & \sim N \left( \text{vec}(\Upsilon), \Sigma \otimes A_{\Upsilon}^{-1} \right) \\
\Sigma & \sim IW(\nu, \Omega)
\end{align*}
$$

Heterogeneous parameters $\theta^1_i$ are thus projected onto demographic variables $Z_i$ using the hyper parameters $\Upsilon$. 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 $\nu = length(\theta^1_i)$ and $W = \nu I$.

## 4 Estimation

Here we describe our Markov Chain Monte Carlo (MCMC) estimation procedure and briefly discuss identification issues.

### 4.1 Joint likelihood

We begin construction of the full model likelihood function with the probability of observed outcomes for a single purchase occasion. We use $l$ to index the $l^{th}$ purchase by customer $i$ in quarter $t$. Using equations (18) and (11), the likelihood for a single purchase occasion may be written as a function of channel expenditures, $\tilde{e}_{itl}$, and category shares, $\tilde{s}_{itl}$:

$$
\mathbb{L}_{itl} = \ell(\tilde{e}_{itl}, \tilde{s}_{itl}) = \ell(\tilde{e}_{itl} = \tilde{e}_{itl}^*) \ell(\tilde{s}_{itl} = \tilde{s}_{itl}^* | \tilde{e}_{itl}^*)
$$

Given the independence across purchase occasions and using equation (21), the likelihood of the collection of purchase occasions for customer $i$ in period $t$ is:

$$
\mathbb{L}_{it} = \ell(L_{it}) \left( \prod_{l=1}^{L_{it}} \mathbb{L}_{itl} \right)_{(L_{it} > 0)}
$$

The likelihood of the $T_i$ period observations for customer $i$ and the total likelihood as a function of all model parameters ($\Theta$) are then given by:

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

### 4.2 Estimation algorithm

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

1. Draw $\theta^1_i$ for each household using a random walk Metropolis-Hastings step with candidate density:

$$
\mathbb{L}_i(\theta^1_i | \theta^2, \gamma, \Sigma, Z_i, \{ e_{itl} \}, \{ s_{itl} \})
$$

2. Draw $\theta^2$ using a random walk Metropolis-Hastings step with candidate density:

$$
\mathbb{L}(\theta^2 | \{ \theta^1_i \}, \gamma, \Sigma, Z_i, \{ e_{itl} \}, \{ s_{itl} \})
$$

3. Draw $\Upsilon | \{ \theta^1_i \}, \Sigma$ from the posterior (assuming conjugate prior)

4. Draw $\Sigma | \{ \theta^1_i \}$ 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 100,000 draws and discard the first 90,000 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 5th draw.

4.2.1 Computational issues

The primary computational challenge for estimation relates to repeatedly solving for the optimal expenditure levels in unchosen channels (including the no-purchase probability in equation (20)). This process requires solving the budget constraint numerically (by observation), which requires repeated evaluation of expected product utility, \(v_c(e_c)\), which in turn requires repeated computation of optimal category shares (for many draws of the channel/category shocks). To overcome the computational burden, we develop an algorithm that: (a) efficiently simulates optimal category shares at given expenditure levels, and (b) accurately approximates the expected product utility as a function of expenditure, using a polynomial interpolant. This approach vastly reduces the number of times expected utility must be simulated when solving the budget constraint equations. For each draw of the parameters, we first obtain a Chebyshev polynomial approximation of \(v_c(e_c)\) for each channel and (customer/quarter) observation. With the (analytically differentiable) polynomial approximations in hand, the budget constraint equations are efficiently solved using a vectorized bisection algorithm.

Our algorithm to compute optimal category shares, which follows Pinjari and Bhat (2010), is explained in detail in Appendix A. The key feature of the algorithm is that it solves for optimal shares via a fixed number of matrix operations, rather than requiring a constrained nonlinear search. For the Chebyshev approximation, we use a 15th order polynomial.\(^{16}\) As the expected utility function is smooth and monotonically increasing in expenditure, these polynomials provide an excellent fit to the simulated values, with absolute deviations averaging less than 0.1%. When simulating \(v_c(e_c)\) and when performing the the Monte Carlo integration over \(\nu\) in equation (11), we use 1000 draws generated from a Halton sequence.

4.3 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 5. 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. 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.

\(^{16}\)We use the extended Chebyshev grid, where interpolation limits are also quadrature nodes. The lower bound of the interpolation domain is zero expenditure. For the upper bound, we use the largest value at which the polynomial will be evaluated, which occurs at the largest quadrature node evaluated in equation (20). Expenditure upper bounds are individual-specific (via budget parameters \(\tau_i\)) and therefore update dynamically during estimation.
consideration ($\mu_i$) is identified by changes in purchase frequency with retail store distance, after controlling for individual-specific retail utility/fixed cost intercepts and CSA time trends (principally, within-customer distance variation due to store entry). We impose a common variance assumption on shopping budgets, which is identified.

In the table, category incidence refers to the frequency of purchases involving the category. Given expenditure levels, individual-specific category incidence rates can identify $\psi_i$, even when categories are purchased discretely (single category baskets) and no price variation is present. Share variation (multiple category baskets) reflects cross-category quantity trade-offs that identify the satiation parameters $\gamma$. Adding price variation allows the variance of category/channel shocks to be identified.\(^{17}\) Since at most $C \times K$ different variance terms are identified and the dual-shock formulation contains $C \times K + 1$ free parameters, we must impose a normalization restriction. We choose to estimate the $\sigma_{ck}$ and normalize the extreme value variance $\sigma_z$ to a sufficiently small value.\(^{18}\) We found that normalizing $\sigma_z = 0.20$ was sufficiently small in our setting. The level of product utility must also be normalized, which we impose by restricting $\psi_1 = 0$.

In the expenditure choice model, the level of (additive) channel utility is normalized by setting online utility ($\omega_1$) to zero. Average retail channel incidence rates given purchase then identify the $\omega_{2i}$. Expenditure variation in the expected utility function $\nu_i(e_i)$ identifies the channel shock variance $\sigma_z$.

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 ($\tau_i$) 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 $f_1 = 7.95$. Under this normalization, the budget level intercept is pinned down by average online expenditure levels, while the retail fixed cost intercept ($\alpha_i$) is identified by the difference in average retail and online expenditure levels. Separate identification of retail utility and retail fixed costs is further aided by our restriction that retail distance enters the cost function but not the channel utility (except indirectly through expenditure levels). We restrict retail fixed costs to be zero at zero distance through the specified functional form (which has flexible scale and curvature). The retail fixed cost distance parameter ($\kappa$) is thus identified by changes in channel choice and expenditure level with retail store distance, after controlling for individual-specific retail utility/fixed cost intercepts and CSA time trends (principally, within-customer distance variation due to store entry). We impose a common variance assumption on shopping budgets, which is identified by aggregate within-subject expenditure variation.

Finally, the consideration arrival rate intercepts $\delta_i$ are identified by customer-specific average purchase frequencies, while the $\beta^\delta$ are identified by common seasonal changes in purchase rates. The effect of distance on consideration ($\zeta$) is identified by changes in purchase frequency with retail store distance, after controlling for individual-specific purchase rates and seasonal effects.

\(^{17}\)To see this, note that in equation (9b), the price term has no coefficient. When shock variances are freely estimated, effectively becomes the price coefficient for category $k$ in channel $c$, since utility is scaled by this variance when entering the likelihood.

\(^{18}\)As a practical matter, this translates to setting $\sigma_z$ small enough to get significant estimates for the $\sigma_{ck}$. See Bhat (2008) for an extended discussion of alternative normalizing procedures.
5 Results

5.1 Model estimates

<table>
<thead>
<tr>
<th>Product utility parameters</th>
<th>heterogeneity</th>
<th>mean</th>
<th>std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satiation</td>
<td>γ1</td>
<td>N</td>
<td>4.1421</td>
</tr>
<tr>
<td>online shock variance</td>
<td>σv11</td>
<td>N</td>
<td>0.6697</td>
</tr>
<tr>
<td>retail shock variance</td>
<td>σv21</td>
<td>N</td>
<td>0.7996</td>
</tr>
<tr>
<td>Category 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satiation</td>
<td>γ2</td>
<td>N</td>
<td>2.8043</td>
</tr>
<tr>
<td>baseline utility intercept</td>
<td>ψ2</td>
<td>Y</td>
<td>-0.9974</td>
</tr>
<tr>
<td>online shock variance</td>
<td>σv12</td>
<td>N</td>
<td>1.2714</td>
</tr>
<tr>
<td>retail shock variance</td>
<td>σv22</td>
<td>N</td>
<td>1.2997</td>
</tr>
<tr>
<td>Category 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satiation</td>
<td>γ3</td>
<td>N</td>
<td>5.1304</td>
</tr>
<tr>
<td>baseline utility intercept</td>
<td>ψ3</td>
<td>Y</td>
<td>-0.3441</td>
</tr>
<tr>
<td>online shock variance</td>
<td>σv13</td>
<td>N</td>
<td>0.9061</td>
</tr>
<tr>
<td>retail shock variance</td>
<td>σv23</td>
<td>N</td>
<td>0.5832</td>
</tr>
<tr>
<td>Category 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satiation</td>
<td>γ4</td>
<td>N</td>
<td>1.2561</td>
</tr>
<tr>
<td>baseline utility intercept</td>
<td>ψ4</td>
<td>Y</td>
<td>-1.0104</td>
</tr>
<tr>
<td>online shock variance</td>
<td>σv14</td>
<td>N</td>
<td>1.9003</td>
</tr>
<tr>
<td>retail shock variance</td>
<td>σv24</td>
<td>N</td>
<td>0.8572</td>
</tr>
<tr>
<td>Category 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satiation</td>
<td>γ5</td>
<td>N</td>
<td>1.4806</td>
</tr>
<tr>
<td>baseline utility intercept</td>
<td>ψ5</td>
<td>Y</td>
<td>-1.1394</td>
</tr>
<tr>
<td>online shock variance</td>
<td>σv15</td>
<td>N</td>
<td>1.5095</td>
</tr>
<tr>
<td>retail shock variance</td>
<td>σv25</td>
<td>N</td>
<td>1.2350</td>
</tr>
<tr>
<td>Category 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satiation</td>
<td>γ6</td>
<td>N</td>
<td>10.2127</td>
</tr>
<tr>
<td>baseline utility intercept</td>
<td>ψ6</td>
<td>Y</td>
<td>-0.9541</td>
</tr>
<tr>
<td>online shock variance</td>
<td>σv16</td>
<td>N</td>
<td>1.3718</td>
</tr>
<tr>
<td>retail shock variance</td>
<td>σv26</td>
<td>N</td>
<td>0.4043</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel utility parameters</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail intercept</td>
<td>ωt</td>
<td>Y</td>
<td>0.8370</td>
</tr>
<tr>
<td>Common channel shock variance</td>
<td>σξ</td>
<td>N</td>
<td>0.8571</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel cost parameters</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail distance scale</td>
<td>α</td>
<td>N</td>
<td>3.1991</td>
</tr>
<tr>
<td>distance power</td>
<td>κ</td>
<td>N</td>
<td>0.6799</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expenditure parameters</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Budget* intercept</td>
<td>τ</td>
<td>Y</td>
<td>4.9282</td>
</tr>
<tr>
<td>shock variance</td>
<td>ση</td>
<td>N</td>
<td>0.5832</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consideration arrival parameters</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>δ</td>
<td>Y</td>
<td>-1.2201</td>
</tr>
<tr>
<td>I(1st quarter)</td>
<td>β1δ</td>
<td>N</td>
<td>-0.0036</td>
</tr>
<tr>
<td>I(3rd quarter)</td>
<td>β2δ</td>
<td>N</td>
<td>0.0012</td>
</tr>
<tr>
<td>I(4th quarter)</td>
<td>β3δ</td>
<td>N</td>
<td>0.0017</td>
</tr>
<tr>
<td>distance power</td>
<td>ζ</td>
<td>N</td>
<td>-0.4651</td>
</tr>
</tbody>
</table>

# obs = 57,554 household x quarter observations, some with multiple purchases.
* CSA time trends included in budget specification (via $\beta'X'$) not reported for brevity.

Table 6: Model estimates
We report posterior means and standard deviations of our parameter estimates in Table 6. For heterogeneous parameters, the reported means average across customers in addition to posterior draws. Among the homogeneous product utility parameters, we note that satiation rates ($\gamma$) vary widely, with category 6 exhibiting the least satiation in quantity (largest $\gamma$) and category 4 exhibiting the most. Also, for 4 categories (all but categories 1 and 2), the online shock variance $\sigma_{\eta_{1k}^2}$ exceeds the retail shock variance $\sigma_{\eta_{2k}^2}$. This result is consistent with the online channel having fewer stock-outs or providing more product information for these categories. Categories 1 and 2 are the ones that we expect have the greatest need to assess physical fit, which suggests why the variance is higher in the retail channel for these categories. Category 6 has a much higher variance in the online channel, which aligns with the firm’s informal guidance that this category has the highest SKU-level stock-out rates in the retail channel.

(a) Category 2 utility ($\psi_2$)  
(b) Category 3 utility ($\psi_3$)  
(c) Category 4 utility ($\psi_4$)  
(d) Category 5 utility ($\psi_5$)  
(e) Category 6 utility ($\psi_6$)  
(f) Retail channel utility ($\omega_2$)  
(g) Budget ($\tau$)  
(h) Retail fixed cost ($\log(\alpha)$)  
(i) Consideration arrival ($\delta$)

Figure 10: Heterogeneous parameter frequency distributions

The sample means of the heterogeneous category intercepts ($\psi_k$) are all negative, implying category 1 (normalized to zero) is the most preferred category, followed by category 3, etc. We summarize the distribution of the heterogeneous category intercepts ($\psi$) and other heterogeneous parameters ($\omega_2$, $\tau$, $\alpha$, $\delta$) in Figure 10. Figure 10 indicates, for example, that valuations are more dispersed for categories 4, 5, and 6 than for categories 2 and 3. There is also evidence of some bimodality in channel and category preferences.

In Table 7, we report the correlation pattern among the heterogeneous parameters. Among the category utility intercepts, we see significant preference correlation among categories 3 and 4 (0.40) as well as categories 4 and 6 (0.34). There is substantial negative correlation between retail channel utility and the utilities for categories 4 and 6 (−0.37, −0.40 respectively), suggesting these categories are preferred by consumers who favor the online channel. Budget levels are strongly negatively correlated with consumers with a preference for category 5 (−0.61). Retail fixed costs are negatively correlated with retail utility (−0.44) and positively correlated with consideration arrival rates (0.32).
Table 7: Correlations among heterogeneous parameters

<table>
<thead>
<tr>
<th></th>
<th>$\psi_2$</th>
<th>$\psi_3$</th>
<th>$\psi_4$</th>
<th>$\psi_5$</th>
<th>$\psi_6$</th>
<th>$\tau$</th>
<th>$\omega_2$</th>
<th>$\alpha$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>category 2 utility</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>category 3 utility</td>
<td>-0.13</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>category 4 utility</td>
<td>0.24</td>
<td>0.40</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>category 5 utility</td>
<td>0.11</td>
<td>0.06</td>
<td>-0.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>category 6 utility</td>
<td>0.16</td>
<td>0.18</td>
<td>0.34</td>
<td>0.25</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>budget $\tau$</td>
<td>-0.10</td>
<td>-0.01</td>
<td>-0.22</td>
<td>-0.61</td>
<td>-0.28</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>retail channel utility $\omega_2$</td>
<td>-0.04</td>
<td>-0.18</td>
<td>-0.37</td>
<td>0.02</td>
<td>-0.40</td>
<td>0.05</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>retail fixed costs $\alpha$</td>
<td>-0.20</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.11</td>
<td>-0.01</td>
<td>0.11</td>
<td>-0.44</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>consideration arrival $\delta$</td>
<td>0.07</td>
<td>-0.03</td>
<td>0.07</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.08</td>
<td>-0.03</td>
<td>0.32</td>
<td>1</td>
</tr>
</tbody>
</table>

In Table 12 of Appendix B we report report estimates of the hyper parameters $\Theta$, 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 a significant relationship between older age and greater retail preference (consistent with (Ansari et al., 2008)), relative preference for category 1, higher budget levels and higher retail transportation costs. Other significant relationships include: i) higher incomes correlate with lower retail fixed costs and lower valuations of categories 4, 5, and 6, ii) larger white populations correlate with lower preference for category 6 and higher preference for the retail channel, iii) more educated populations correlate with lower retail fixed costs and consideration arrival rates, and iv) longer mean commuting times correspond with higher retail transportation costs and lower consideration arrival rates and category 6 valuations.

Finally we turn to distance-related effects. The consideration arrival parameter $\zeta$ captures the effect of retail distance on brand consideration. The negative (highly significant) coefficient for $\zeta$ implies that as distance decreases, consideration arrival rates (and hence purchase frequencies) will go up. At the same time, the magnitude of retail fixed costs decreases in accordance with the estimated cost function, which we plot in Figure 11 (where retail fixed costs are plotted using the sample average scale effect $\alpha = \bar{\alpha}$). The retail cost function is concave ($\kappa < 1$) such that the marginal cost of traveling to the store declines with distance. On a purely fixed cost basis, the average consumer would be indifferent between the channels at a distance of approximately 8 miles from the store. However, fixed costs are only one consideration for consumers: individual category/channel preferences and product prices also factor into channel choice trade-offs. In Sections 5.2 and 5.3, we simulate from the model to measure retail distance effects on quarterly channel expenditure patterns.  

For example, transportation costs may decline in distance due to greater use of highways (and hence better fuel milage) when traveling longer distances. Such effects are magnified to the extent that estimated fixed costs capture the opportunity cost of time spent traveling.
5.2 Price and retail distance elasticities

We derive price and distance elasticity measures from the model estimates by computing (changes in) expected channel expenditures and category shares. When computing expectations, we aggregate over purchase occasions in the quarter. For example, expected quarterly channel expenditures ($\bar{e}_c$) are calculated for each customer/quarter observation using:

$$\bar{e}_c = E \left[ \sum_{t=1}^{T} e_{c,t}^* \right] = \mu E [e_c^*].$$

For unit demand price elasticities, we work with category/channel quantity indices, $q_{ck} = \mu E \left[ q_{ck}^* \right]$. For all outcome variables of interest, we compute the partial derivatives with respect to prices by observation (either quarter or purchase occasion) and then aggregate over observations. We then calculate the price elasticities by multiplying by price and dividing by the aggregated outcome measure.

The aggregate demand elasticities are reported in the first two columns of Table 8. Unit demand is more responsive to price in the retail channel for categories 3 and 6, while demand for other categories is more responsive to online channel price changes. Cross-channel differences are largest (in relative terms) for category 1, the only category which provides significantly more information in the retail channel (as indicated by the shock variance). In the last two columns of Table 8, we report total (online + retail) expenditure price elasticities. The total expenditures effectively capture the net effect of price changes on substitution with the outside good. That the last two columns are inelastic as compared to the first two reflects that most price-induced switching occurs among the inside categories.

<table>
<thead>
<tr>
<th>category</th>
<th>demand (quantity) online</th>
<th>demand (quantity) retail</th>
<th>total expenditures online</th>
<th>total expenditures retail</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2.113</td>
<td>-0.796</td>
<td>-0.011</td>
<td>-0.034</td>
</tr>
<tr>
<td>2</td>
<td>-1.830</td>
<td>-1.378</td>
<td>-0.005</td>
<td>-0.023</td>
</tr>
<tr>
<td>3</td>
<td>-1.535</td>
<td>-2.167</td>
<td>-0.004</td>
<td>-0.003</td>
</tr>
<tr>
<td>4</td>
<td>-2.347</td>
<td>-1.261</td>
<td>-0.031</td>
<td>-0.028</td>
</tr>
<tr>
<td>5</td>
<td>-1.874</td>
<td>-1.215</td>
<td>-0.012</td>
<td>-0.016</td>
</tr>
<tr>
<td>6</td>
<td>-2.700</td>
<td>-3.909</td>
<td>-0.009</td>
<td>-0.000</td>
</tr>
</tbody>
</table>

Table 8: Elasticities of category demand (quantity) and total expenditures with respect to price

Of particular interest are retail distance effects on total expenditures and expenditures by channel. In Table 9 we report expected quarterly channel expenditure and purchase frequency elasticities with respect to retail store distance. 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 1.97%. The
total expenditure increase arises primarily due to an increase in the total purchase frequency (1.95%), which we attribute to increased brand consideration from the retail presence. In addition, lower retail transportation costs increase the probability of a retail purchase such that the number of retail purchases increases 3.20% while the number of online purchases declines 0.03%. The recapture of consumer transportation cost savings directly increases the expenditure per retail purchase, as reflected by the fact that retail expenditures rise more rapidly (3.35%) than the number of retail purchases (3.20%). A further indirect effect of lower retail transportation costs is that online expenditures per purchase rise – consumers shift lower budget arrival events to the retail channel, leaving fewer but higher budget arrival events allocated to the online channel. This pattern is reflected by the fact that quarterly online expenditures rise by 0.07% while the number of online purchases declines 0.03%.

<table>
<thead>
<tr>
<th>outcome</th>
<th>elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>total expenditures</td>
<td>-0.197</td>
</tr>
<tr>
<td>online expenditures</td>
<td>-0.007</td>
</tr>
<tr>
<td>retail expenditures</td>
<td>-0.335</td>
</tr>
<tr>
<td>total purchase frequency</td>
<td>-0.195</td>
</tr>
<tr>
<td>online purchase frequency</td>
<td>0.003</td>
</tr>
<tr>
<td>retail purchase frequency</td>
<td>-0.320</td>
</tr>
</tbody>
</table>

Table 9: Elasticities with respect to retail store distance

5.3 Channel revenue retail distance dependence

Our distance elasticities reported in Section 5.2 measure the average expenditure response to changes in retail distance, evaluated at the empirical distribution of customer/store locations. 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 12, we plot these expected quarterly revenues by channel and total quarterly expected revenue versus retail distance. In the right hand panel of Figure 12, we plot point elasticities corresponding to the curves in the left hand panel.

In the left panel of Figure 12, total revenues increase monotonically as distance decreases, from approximately $47 per customer per quarter at 100 miles retail distance to $59 at ten miles and $80 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. Preferences for the online channel are sufficiently high among some customers that online revenues remain appreciable even as retail distance approaches zero miles.

Retail and online channels contribute equally to firm revenues at a distance of approximately 39 miles from the store. Given our previous finding that consumer costs equate at approximately 8 miles, the fact that online revenues are expected to be about 36% of total revenues at this distance demonstrates the added utility many consumers get from shopping in the retail channel. Outside this radius, the curvature of the revenue functions is moderate and well approximated by a linear function. Linearizing the curves in this range and comparing slopes reveals that as distance decreases, retail revenues increase at approximately 2.5 times the rate that online revenues decline. In terms of elasticities (right panel of Figure 12), we see that distance elasticities for retail revenues are always negative and decline monotonically with distance. In contrast to what we would expect if only channel cannibalization effects were present, online revenue elasticities are not always positive and the relationship with distance is not monotonic.
We note that in the left panel of Figure 12, online revenues reach a minimum at approximately five miles from retail stores. This distance represents the point at which the market-expanding consideration effect stops dominating the online channel cannibalization effect (via lower retail fixed costs) for the “average” customer. The distribution of these distances across consumers is shown in Figure 13. The considerable heterogeneity across consumers underscores the importance of understanding demand in the very local region around potential entry points if the brand consideration effect of retail locations is to be leveraged by managers.

We emphasize that the preceding results reflect demand among existing customers. In our retail location counterfactual (Section 6.1), we expand our analysis to include new customer acquisition, including its dependence on retail distance.

5.4 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 C, we provide distribution plots for: a) the number of simulated purchase

Figure 12: Expected quarterly revenue by channel with equidistant customers, as a function of retail store distance

Figure 13: Distribution (across customers) of distances where online revenues increase with retail proximity
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. The distribution of simulated outcomes closely replicates the distribution of observed outcomes.

6 Applications

In this section, we demonstrate the application of our results to the selection of retail entry locations and to exploring counterfactual channel pricing policies.

6.1 Retail location selection

The choice of where to locate retail stores is a difficult problem that many firms face. In this section, we demonstrate how our model may inform such decisions. We abstract from cost concerns by assuming the firm’s objective is to maximize its short-run (5 year/20 quarter) expected revenue. It would be trivial to add cost information were it available to us. Similar to the grid approach in Duan and Mela (2009), our method is to generate a comprehensive set of potential entry locations and to calculate the expected incremental revenue generated from locating a store at each of these locations. To maintain high spatial resolution yet keep the exercise tractable, we discretize the set of potential entry locations as Census 2010 block group centroids in the continental United States, and thus evaluate approximately 216,000 candidate locations.

For each candidate location, we compute a revenue index (defined as $\Gamma$) using a two step process. In the first step, we determine the set of markets (other block groups) that would be impacted by entry at the candidate location. This set is generated by identifying block groups for which the distance to the nearest retail outlet would be reduced as a result of entry at the candidate location. In the second step, for each market in the impacted set, we forecast demand under two conditions: first assuming no entry were to occur, and then assuming it does occur. Taking the difference in these forecasts generates a prediction of the incremental contribution of entry.

Our demand forecasts hold prices and existing store locations fixed at their final period values and incorporate revenue contributions from both new and existing customers. We extend our framework to include new customer acquisition by assuming that: i) an exogenous market-level Poisson process drives new customer arrivals, and ii) the preferences of new customers are drawn from the (market specific) heterogeneity distribution described in Section 3.6. In Appendix D, we describe how we build upon the results of Section 2.4.4 and incorporate demographic variables into our prediction of new customer arrivals. Let $I_{mt}$ be the set of existing and simulated new customers in market $m$ and simulated future period $t$, and let $\bar{e}_{itc}$ be the expected expenditure for customer $i$ in period $t$ and channel $c$. Our revenue index is then given by: $\Gamma_m = \sum_{t=1}^{20} \sum_{c=1}^{20} \sum_{i \in I_{mt}} \bar{e}_{itc}$.

Experiment results Since neighboring block groups within a metro region often yield similar revenue indices, it is difficult to discern the relative desirability of broad regions (e.g., cities) using an ungrouped ranking of potential entry locations. We thus summarize the results of the experiment at the MSA (Metropolitan Statistical Area) level. To do this, we assign a MSA-level index using the maximal revenue index among the block groups in that MSA. These indices are reported in Table 10.
### Table 10: Top 10 model-predicted entry locations (MSA level)

Comparing the recommended entry MSAs in Table 10 to the brand’s store locations operative during our study reveals that none of these MSAs have existing retail stores. The experiment thus intuitively suggests that the most desirable locations may be found in markets with no existing brand presence. Lending some face validity to our experiment, we note that since the end of the observation window, the brand has opened store locations in six locations identified in Table 10. To demonstrate that the counterfactual results can help inform both local and wide-geography choices, we generate revenue heat maps for the top ranked MSA, Minneapolis-St. Paul, which are provided in Figure 14. The map identifies the most desirable location by the darkest shade of red, as well as close substitutes in the event entry is not possible in that specific location (due, for example, to zoning restrictions or lack of available rental space).

<table>
<thead>
<tr>
<th>Rank</th>
<th>MSA</th>
<th>Revenue index (SMM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Minneapolis-St. Paul, MN-WI</td>
<td>32.59</td>
</tr>
<tr>
<td>2</td>
<td>Miami-Fort Lauderdale, FL</td>
<td>26.76</td>
</tr>
<tr>
<td>3</td>
<td>Cincinnati-Hamilton, OH-KY-IN</td>
<td>20.59</td>
</tr>
<tr>
<td>4</td>
<td>West Palm Beach-Boca Raton, FL</td>
<td>19.86</td>
</tr>
<tr>
<td>5</td>
<td>Dayton-Springfield, OH</td>
<td>17.27</td>
</tr>
<tr>
<td>6</td>
<td>St. Louis, MO-IL</td>
<td>15.65</td>
</tr>
<tr>
<td>7</td>
<td>Columbus, OH</td>
<td>14.47</td>
</tr>
<tr>
<td>8</td>
<td>Indianapolis, IN</td>
<td>13.05</td>
</tr>
<tr>
<td>9</td>
<td>Louisville, KY-IN</td>
<td>9.41</td>
</tr>
<tr>
<td>10</td>
<td>Salt Lake City-Ogden, UT</td>
<td>7.84</td>
</tr>
</tbody>
</table>

Figure 14: Revenue generation index by Census block group, Minneapolis-St. Paul MSA

### 6.2 Eliminating online shipping costs

Another important issue for multi-channel firms is how to design channel-specific pricing policies. Our model can also be used to assess the demand (and with cost information, profit) implications of various pricing policies. 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 this application, we consider the impact of eliminating online shipping fees, i.e., 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 11 below. For the baseline ($f_1 = 7.95$) and no shipping ($f_1 = 0$)
scenarios, we report quarterly average revenues, number of purchases, and average revenue per purchase for the online channel, retail channel, and both channels combined (total). We then report the change in these measures going from the baseline to no shipping scenarios, in absolute (Δ) and percentage (Δ%) terms.

Eliminating shipping fees increases total revenue by 2.8%, driven by a 1.4% increase in overall purchase frequency and a 1.4% increase in expenditures per purchase. Broken out by channel, online revenues increase by 14.6% while retail revenues decrease by 5.9%. These changes in channel revenues are principally due to switching in purchase incidence from retail to online: the number of online purchases increases by 15.1% while the number of retail purchases declines by 7.1%. By contrast, the effect on revenues per purchase is more modest, with online revenues per purchase declining slightly (-0.5%) and retail revenues per purchase increasing by 1.3%.

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 $1.62 per customer per quarter, implying an expected net loss of $0.03 per customer per quarter. 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.

<table>
<thead>
<tr>
<th>Total (Online + Retail)</th>
<th>Online</th>
<th>Retail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quarterly revenue</td>
<td>Quarterly revenue</td>
<td>Quarterly revenue</td>
</tr>
<tr>
<td>quarterly purchases</td>
<td>quarterly purchases</td>
<td>quarterly purchases</td>
</tr>
<tr>
<td>Rev per purchase</td>
<td>Rev per purchase</td>
<td>Rev per purchase</td>
</tr>
<tr>
<td>baseline</td>
<td>58.360</td>
<td>24.646</td>
</tr>
<tr>
<td>no shipping</td>
<td>59.981</td>
<td>28.240</td>
</tr>
<tr>
<td>Δ</td>
<td>1.621</td>
<td>3.594</td>
</tr>
<tr>
<td>Δ%</td>
<td>2.78%</td>
<td>14.58%</td>
</tr>
</tbody>
</table>

Table 11: No online shipping costs counterfactual results

7 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, our structural model estimates and descriptive regressions provide evidence of a promotional effect of retail proximity – i.e., retail presence raises brand consideration, leading to increased purchase frequency. We document evidence of similar distance-related consideration effects on new customer acquisition. Among existing customers, our structural estimates imply a 10% reduction in retail store distance increases total quarterly revenues by 1.97%, a result of increased brand consideration and partial recapture of consumer transportation cost savings. Online and retail revenues both increase for existing consumers in the near proximity of new retail outlets (within a distance of five miles for the modal consumer).

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 expenditure decisions maximize the expected (second stage, category share optimal) utility subject to a stochastic shopping budget constraint, which includes
both fixed and variable channel transaction costs. Finally, we develop a computationally efficient algorithm to jointly estimate the multi-stage model. While our application is to apparel categories, our model and estimation methodology can be applied in any multi-channel context where the analyst has access to historical customer transaction data.

We also contribute to managerial practice by providing a tool to address challenging problems such as where to locate new retail stores and how to optimize product prices in a multi-channel environment. Our retail entry experiment demonstrates the computational feasibility of exhaustively exploring potential entry locations at high spatial resolution, a process that could be applied iteratively to obtain even more precise predictions of optimal locations. We further demonstrate our model’s utility for evaluating counterfactual channel pricing policies, such as alternative shipping fees.

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. Those with panel data containing high purchase frequencies or long time dimensions could tackle dynamic considerations such as state dependence in channel choices or learning about product categories over time. Relatedly, the role of channel interactions in the consumer’s search process could be explored. While we believe that our focus on apparel categories to some extent limits the scope for dynamics in our analysis (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


Appendices

A Computation of optimal shares and expected utilities

We first demonstrate the calculation of optimal shares when the set of chosen categories is known and then describe the procedure to determine the set of chosen categories. For a given draw of the category shocks \((\varepsilon, \eta)\), without loss of generality assume the first \(M\) categories are chosen. For chosen categories, the Kuhn-Tucker condition \(\frac{\partial L}{\partial s_k} = 0\) implies \(\frac{\Psi_{ck}\varepsilon^*_c}{p_c} = \lambda\), which may be solved for the optimal share: \(s^*_k = \frac{\Psi_{ck}}{\lambda} - \frac{\eta p_c}{\varepsilon^*_c}\). Next, we use the budget constraint equation to eliminate the Lagrange multiplier \(\lambda\): \(1 = \sum_{k=1}^{M} s^*_k = \sum_{k=1}^{M} \left(\frac{\Psi_{ck}}{\lambda} - \frac{\eta p_c}{\varepsilon^*_c}\right)\).

Solving for \(\lambda\) gives \(\lambda = \frac{e^*_c}{\sum_{k=1}^{M} \eta p_c}\). Then, substituting this expression for \(\lambda\) back into the optimal share equation above gives:

\[
s^*_k = \frac{\Psi_{ck}}{e^*_c} \left(\frac{e^*_c + \sum_{k=1}^{M} \eta p_c}{\sum_{k=1}^{M} \Psi_{ck}} - \frac{p_c}{\lambda}\right)
\]

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 \(\frac{\Psi_{ck}}{\bar{p}_{ck}}\) 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 \((\varepsilon, \eta)\)

2. Compute the price normalized baseline utilities for all \(K\) categories: \(\bar{Y}_k = \frac{\Psi_{ck}}{\bar{p}_{ck}}\)

3. Sort the categories from highest to lowest \(Y\); denote quantities sorted in this order with a tilde, e.g. \(\tilde{Y}\)

4. Iteratively compute the Lagrange multiplier, assuming the first \(m\) categories are chosen: \(\lambda^m = \frac{e^*_c m\sum_{j=1}^{m} \eta \beta_{cj}}{e^*_c + m\sum_{j=1}^{m} \eta \beta_{cj}}\)

5. Determine the number of chosen categories by the relation: \(M = \sum_{m=1}^{K} l(\tilde{Y}_m \geq \lambda^m)\)

6. Compute the optimal shares for the chosen categories as \(s^*_m = \frac{\tilde{Y}_m}{e^*_c} \left(\frac{e^*_c + \sum_{k=1}^{m} \bar{p}_{cm}}{\sum_{m=1}^{M} \Psi_{cm}} - \bar{p}_{cm}\right)\) for \(m \leq M\) and \(s^*_m = 0\) for \(m > M\)

7. Invert the sort order to restore the original category ordering, yielding \(s^*_k\)
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).

Once optimal shares are in hand, it is a simple matter to evaluate the expected (indirect) utility by substituting optimal shares into equation (7d).

### B Hyper parameter results

<table>
<thead>
<tr>
<th>Hyper-parameters (T)</th>
<th>Intercept</th>
<th>Age</th>
<th>White</th>
<th>College</th>
<th>HH Income</th>
<th>Rural</th>
<th>Commute</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Product utility parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category 2 utility intercept $\psi_2$</td>
<td>-0.6219</td>
<td>-0.0075</td>
<td>0.0573</td>
<td>0.1070</td>
<td>-0.0006</td>
<td>-0.0145</td>
<td>-0.0040</td>
</tr>
<tr>
<td></td>
<td>(0.1976)</td>
<td>(0.0021)</td>
<td>(0.1717)</td>
<td>(0.1395)</td>
<td>(0.0006)</td>
<td>(0.0868)</td>
<td>(0.0030)</td>
</tr>
<tr>
<td>Category 3 utility intercept $\psi_3$</td>
<td>-0.0454</td>
<td>-0.0066</td>
<td>-0.0455</td>
<td>0.0688</td>
<td>-0.0005</td>
<td>-0.0236</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.1960)</td>
<td>(0.0021)</td>
<td>(0.1345)</td>
<td>(0.1351)</td>
<td>(0.0006)</td>
<td>(0.0838)</td>
<td>(0.0036)</td>
</tr>
<tr>
<td>Category 4 utility intercept $\psi_4$</td>
<td>-0.3650</td>
<td>-0.0135</td>
<td>0.0296</td>
<td>0.1407</td>
<td>-0.0013</td>
<td>-0.0154</td>
<td>-0.0025</td>
</tr>
<tr>
<td></td>
<td>(0.1801)</td>
<td>(0.0018)</td>
<td>(0.1224)</td>
<td>(0.1457)</td>
<td>(0.0005)</td>
<td>(0.0850)</td>
<td>(0.0035)</td>
</tr>
<tr>
<td>Category 5 utility intercept $\psi_5$</td>
<td>-0.4244</td>
<td>-0.0074</td>
<td>-0.1832</td>
<td>0.0599</td>
<td>-0.0023</td>
<td>-0.0543</td>
<td>-0.0027</td>
</tr>
<tr>
<td></td>
<td>(0.1624)</td>
<td>(0.0019)</td>
<td>(0.1482)</td>
<td>(0.1393)</td>
<td>(0.0006)</td>
<td>(0.0861)</td>
<td>(0.0031)</td>
</tr>
<tr>
<td>Category 6 utility intercept $\psi_6$</td>
<td>0.4065</td>
<td>-0.0125</td>
<td>-0.4731</td>
<td>-0.0161</td>
<td>-0.0021</td>
<td>-0.0383</td>
<td>-0.0093</td>
</tr>
<tr>
<td></td>
<td>(0.1418)</td>
<td>(0.0019)</td>
<td>(0.1025)</td>
<td>(0.1081)</td>
<td>(0.0005)</td>
<td>(0.0661)</td>
<td>(0.0028)</td>
</tr>
<tr>
<td><strong>Channel utility parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retail utility intercept $\omega_2$</td>
<td>-0.0140</td>
<td>0.0117</td>
<td>0.4736</td>
<td>0.0730</td>
<td>-0.0005</td>
<td>0.0906</td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.1766)</td>
<td>(0.0015)</td>
<td>(0.1227)</td>
<td>(0.1662)</td>
<td>(0.0007)</td>
<td>(0.0635)</td>
<td>(0.0033)</td>
</tr>
<tr>
<td>Budget intercept $\tau$</td>
<td>4.3337</td>
<td>0.0089</td>
<td>0.1232</td>
<td>0.0188</td>
<td>0.0009</td>
<td>0.0434</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>(0.1681)</td>
<td>(0.0018)</td>
<td>(0.1246)</td>
<td>(0.1335)</td>
<td>(0.0006)</td>
<td>(0.0791)</td>
<td>(0.0034)</td>
</tr>
<tr>
<td><strong>Channel cost parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>distance scale $\alpha$</td>
<td>-0.3469</td>
<td>0.0156</td>
<td>-0.0927</td>
<td>-0.3700</td>
<td>-0.0020</td>
<td>-0.0519</td>
<td>0.0311</td>
</tr>
<tr>
<td></td>
<td>(0.1234)</td>
<td>(0.0014)</td>
<td>(0.1242)</td>
<td>(0.1116)</td>
<td>(0.0007)</td>
<td>(0.0856)</td>
<td>(0.0030)</td>
</tr>
<tr>
<td><strong>Consideration arrival parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>intercept $\delta$</td>
<td>-0.8813</td>
<td>-0.0018</td>
<td>0.0339</td>
<td>-0.4179</td>
<td>0.0004</td>
<td>0.0690</td>
<td>-0.0041</td>
</tr>
<tr>
<td></td>
<td>(0.1061)</td>
<td>(0.0013)</td>
<td>(0.0979)</td>
<td>(0.0993)</td>
<td>(0.0004)</td>
<td>(0.0589)</td>
<td>(0.0020)</td>
</tr>
</tbody>
</table>

Table 12: Model hyper parameter estimates
C Model fit exhibits

![Figure 15: Simulated frequency of purchase occasions per quarter](image)

<table>
<thead>
<tr>
<th>moments</th>
<th>data</th>
<th>simulation</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>purchases (#/qtr)</td>
<td>0.51</td>
<td>0.51</td>
<td>0.00%</td>
</tr>
<tr>
<td>online expenditures ($/qtr)</td>
<td>30.22</td>
<td>32.19</td>
<td>+6.52%</td>
</tr>
<tr>
<td>retail expenditures ($/qtr)</td>
<td>33.99</td>
<td>35.48</td>
<td>+4.38%</td>
</tr>
<tr>
<td>online share of expenditures (per qtr)</td>
<td>47.06%</td>
<td>47.57%</td>
<td>+1.08%</td>
</tr>
<tr>
<td>online share of purchases</td>
<td>43.31%</td>
<td>43.42%</td>
<td>+0.25%</td>
</tr>
<tr>
<td>online expenditure/purchase ($)</td>
<td>138.03</td>
<td>146.64</td>
<td>+6.24%</td>
</tr>
<tr>
<td>retail expenditure/purchase ($)</td>
<td>118.59</td>
<td>124.03</td>
<td>+4.61%</td>
</tr>
</tbody>
</table>

Table 13: Data vs. Simulation Moments

![Figure 16: Simulated expenditure distributions by channel](image)

![Figure 17: Simulated average category shares by channel](image)

D Demographic effects on new customer acquisition

For the purpose of evaluating our entry counterfactual, it is useful to ascertain the influence of market demographics on new customer acquisition because the counterfactual involves simulating acquisition rates for markets not observed in the data. To obtain estimates, we extend the results of the Poisson fixed effects regression in Section...
2.4.4. To obtain these estimates without weakening the controls identifying the distance effect ($\alpha$), we perform a separate Poisson regression without market fixed effects, but where we constrain $\alpha$ to be as estimated in Section 2.4.4. To that specification, we add demographic information obtained from Census 2010 block group records ($z_m$) and period fixed effects ($\mu_t$):

\[ N_{mt} \sim \text{Poisson}(\rho_{mt}), \quad \log(\rho_{mt}) = \alpha \log(d_{mt}) + \beta z_m + \mu_t + \log(\text{pop}_m - E_{mt}) \] (25)

The $\beta$ parameter estimates are reported in Table 14 below – all effects are significant at the 1% level. The results suggest strong positive correlation with market average age, median income, fraction of population white and fraction of population with college degrees. There is a negative correlation associated with rural markets.

<table>
<thead>
<tr>
<th>regressor</th>
<th>estimate</th>
<th>std err</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>0.0187</td>
<td>0.0015</td>
</tr>
<tr>
<td>income</td>
<td>0.0015</td>
<td>0.0003</td>
</tr>
<tr>
<td>rural</td>
<td>-0.2036</td>
<td>0.0435</td>
</tr>
<tr>
<td>white</td>
<td>0.7387</td>
<td>0.0670</td>
</tr>
<tr>
<td>college</td>
<td>3.4792</td>
<td>0.0655</td>
</tr>
<tr>
<td>commute</td>
<td>1.0404</td>
<td>0.2032</td>
</tr>
<tr>
<td>quarter</td>
<td>-0.0047</td>
<td>0.0017</td>
</tr>
<tr>
<td>I(1st quarter)</td>
<td>-0.0093</td>
<td>0.0057</td>
</tr>
<tr>
<td>I(3rd quarter)</td>
<td>0.0583</td>
<td>0.0372</td>
</tr>
<tr>
<td>I(4th quarter)</td>
<td>0.2434</td>
<td>0.0335</td>
</tr>
<tr>
<td>constant</td>
<td>-14.6260</td>
<td>0.1444</td>
</tr>
</tbody>
</table>

Table 14: Estimates of demographic effects on new customer acquisition
Technical Appendix - Not Intended for Publication

E   Details of the category share model

E.1 Kuhn-Tucker conditions

As stated in the text, the Kuhn-Tucker conditions are: a) \( \frac{\partial \mathcal{L}}{\partial s_k} \leq 0 \) (stationarity), b) \( s_k \geq 0 \) (primal feasibility), and c) \( \frac{\partial \mathcal{L}}{\partial s_k} s_k = 0 \) (complementary slackness), while the Lagrangian is given by: 

\[
\mathcal{L} = \sum_{k=1}^{K} \Psi_{ck} \gamma_k \log \left( \frac{s_k e^c}{p_{ck} \gamma_k} + 1 \right) - \lambda \left( \sum_{k=1}^{K} s_k - 1 \right).
\]

Recall that without loss of generality, we assume the first category is chosen (categories may be re-labeled to assure this is the case) and thus \( s_1^* > 0 \). By condition (c) \( \frac{\partial \mathcal{L}}{\partial s_1} = 0 \), the Lagrange multiplier is given by \( \lambda = \frac{c^\gamma \Psi_{c1}}{\lambda_c \left( \frac{\gamma_{c1}}{\gamma_{c1}} + 1 \right)} \). For other chosen categories (where \( s_k^* > 0 \)), we must have \( \frac{\partial \mathcal{L}}{\partial s_k} = \frac{c^\gamma \Psi_{ck}}{\lambda_c \left( \frac{\gamma_{ck}}{\gamma_{ck}} + 1 \right)} - \lambda = 0 \). Rearranging, taking logarithms, and substituting in from Table 4 yields:

\[
\log (\Psi_{ck}) - \log \left( \frac{s_k e^c}{p_{ck} \gamma_k} + 1 \right) - \log (p_{ck}) = \log (\Psi_{c1}) - \log \left( \frac{s_1 e^c}{p_{c1} \gamma_1} + 1 \right) - \log (p_{c1})
\]

\[
\psi_k + \eta_k - \log \left( \frac{s_k e^c}{p_{ck} \gamma_k} + 1 \right) - \log (p_{ck}) + \epsilon_{ck} = \psi_1 + \eta_1 - \log \left( \frac{s_1 e^c}{p_{c1} \gamma_1} + 1 \right) - \log (p_{c1}) + \epsilon_{c1}
\]

\[
g_{ck} + \epsilon_{ck} = g_{c1} + \epsilon_{c1}
\]

Where \( g_{ck} = \psi_k + \eta_k - \log \left( \frac{s_k e^c}{p_{ck} \gamma_k} + 1 \right) - \log (p_{ck}) \). For non-chosen categories, a similar derivation yields \( g_{ck} + \epsilon_{ck} < g_{c1} + \epsilon_{c1} \). These conditions correspond to those stated in equation (9a) in the text.

E.2 Likelihood function

This derivation follows Bhat (2008). We begin by evaluating the integrals with respect to the \( \epsilon_k \) terms in equation (10):

\[
\ell(\delta = \delta^c | e^c) = \left| J_{F^c \rightarrow \delta^c} \right| \int_{\eta}^{+\infty} \int_{\epsilon_1 = -\infty}^{+\infty} \cdots \int_{\epsilon_{M+1} = -\infty}^{+\infty} \frac{f(\overline{\epsilon})(f(\overline{\eta})) d\overline{\epsilon} d\overline{\eta}}{f(\overline{\epsilon}) f(\overline{\eta})}
\]

\[
= \left| J_{F^c \rightarrow \delta^c} \right| \int_{\epsilon_1 = -\infty}^{+\infty} \left\{ \prod_{j=2}^{M+1} \frac{1}{\sigma_e} \lambda \left[ \frac{g_{c1} - g_{cj} + \epsilon_1}{\sigma_e} \right] \right\}
\]

\[
\times \left\{ \prod_{j=M+1}^{K} \Lambda \left[ \frac{g_{c1} - g_{cj} + \epsilon_1}{\sigma_e} \right] \right\} \frac{1}{\sigma_e} \lambda \left[ \frac{\epsilon_1}{\sigma_e} \right] d\epsilon_1 f(\overline{\eta}) d\overline{\eta}
\]

where \( \lambda(x) = e^{-x+x^{-1}} \) and \( \Lambda(x) = e^{-e^{-x}} \) are the standard extreme value pdf and cdf, respectively. Substi-
The Jacobian determinant is then:

\[
\ell(\tilde{s} = \tilde{s}^b | e_c^*) = \left| J_{\tilde{\eta}M \to \tilde{\eta}M} \right| \frac{1}{\sigma_c^{M-1}} \prod_{j=2}^{M} \exp \left( \frac{g_{c1} - g_{cj}}{\sigma_c} \right) \times \int_{-\infty}^{+\infty} e^{-\sum_{j=1}^{K} \exp \left( -\left( \frac{g_{c1} - g_{cj}}{\sigma_c} \right) \right)} \left( \frac{1}{\sigma_c} e^{-\frac{\epsilon_1}{\sigma_c}} \right)^{M-1} \frac{1}{\sigma_c} e^{-\frac{\epsilon_1}{\sigma_c}} d\epsilon_1 f(\tilde{\eta}) d\tilde{\eta}
\]

Let \( u = e^{-\frac{\epsilon_1}{\sigma_c}} \). Then, the inner integral in the second line above may be written:

\[
\int_{-\infty}^{+\infty} e^{-\sum_{j=1}^{K} \exp \left( -\left( \frac{g_{c1} - g_{cj}}{\sigma_c} \right) \right)} \left( \frac{1}{\sigma_c} e^{-\frac{\epsilon_1}{\sigma_c}} \right)^{M-1} \frac{1}{\sigma_c} e^{-\frac{\epsilon_1}{\sigma_c}} d\epsilon_1 = \int_{u=0}^{+\infty} e^{-u \sum_{j=1}^{K} \exp \left( -\left( \frac{g_{c1} - g_{cj}}{\sigma_c} \right) \right)} u^{M-1} du
\]

\[
= (M-1)! \left( \sum_{j=1}^{K} \exp \left( -\left( \frac{g_{c1} - g_{cj}}{\sigma_c} \right) \right) \right)^{-M}
\]

Substituting back in and simplifying yields:

\[
\ell(\tilde{s} = \tilde{s}^b | e_c^*) = \left| J_{\tilde{\eta}M \to \tilde{\eta}M} \right| \frac{(M-1)!}{\sigma_c^{M-1}} \prod_{j=2}^{M} \exp \left( \frac{g_{c1} - g_{cj}}{\sigma_c} \right) \left( \sum_{j=1}^{K} \exp \left( -\left( \frac{g_{c1} - g_{cj}}{\sigma_c} \right) \right) \right)^{-M} f(\tilde{\eta}) d\tilde{\eta}
\]

Next, consider the elements of the Jacobian:

\[
J_{ij} = \frac{\partial \tilde{\epsilon}_{i+1}}{\partial s_{j+1}} = \frac{\partial}{\partial s_{j+1}} (g_{c1} - g_{c,j+1} + \epsilon_1) \quad \forall i, j = 1, ..., M - 1
\]

\[
= \frac{\partial}{\partial s_{j+1}} \left( \psi_1 + \eta_{c1} - \log \left( \frac{s_{i+1}^c e_c^*}{p_c + 1} \right) - \log (p_{c1}) - \left( \psi_{i+1} + \eta_{c,i+1} - \log \left( \frac{s_{i+1}^c e_c^*}{p_{c,i+1} + 1} \right) - \log (p_{c,i+1}) \right) + \epsilon_1 \right)
\]

\[
= \frac{\partial}{\partial s_{j+1}} \left( -\log \left( \frac{b - e_c^* \sum_{k=2}^{K} s_{k}^c}{p_{c1} + 1} + 1 \right) + \log \left( \frac{s_{i+1}^c e_c^*}{p_{c,i+1} + 1} + 1 \right) \right)
\]

\[
= \frac{1}{b - e_c^* \sum_{k=2}^{K} s_{k}^c} + \frac{1}{p_{c1} + 1} + \frac{1}{p_{c,i+1} + 1} I(i = j)
\]

\[
= \frac{1}{s_{i+1}^c e_c^* + p_{c1} + 1} + \frac{1}{s_{i+1}^c e_c^* + p_{c,i+1} + 1} I(i = j)
\]

The Jacobian determinant is then:
\[ |J_{EM \rightarrow SM}| = \begin{vmatrix} \frac{1}{s_1 e_j^* + p_c y_j} & \frac{1}{s_2 e_j^* + p_c y_j} & \cdots & \frac{1}{s_M e_j^* + p_c y_j} \\ \frac{1}{s_1 e_j^* + p_c y_j} + \frac{1}{s_2 e_j^* + p_c y_j} & \frac{1}{s_3 e_j^* + p_c y_j} & \cdots & \frac{1}{s_M e_j^* + p_c y_j} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{s_1 e_j^* + p_c y_j} + \frac{1}{s_2 e_j^* + p_c y_j} + \cdots + \frac{1}{s_M e_j^* + p_c y_j} & \frac{1}{s_M e_j^* + p_c y_j} + \frac{1}{s_{M+1} e_j^* + p_c y_j} & \cdots & \frac{1}{s_{M+M} e_j^* + p_c y_j} \end{vmatrix} \]

Substituting this expression for the Jacobian into equation (26) above gives the fully simplified likelihood:

\[
\ell(\vec{s} = \vec{s}^\ast | e_c^\ast) = \frac{(M - 1)!}{\sigma_e^{M-1}} \left( \prod_{j=1}^{M} \frac{1}{s_j e_c^\ast + p_c y_j^j} \right) \left( \sum_{j=1}^{M} s_j^j e_c^\ast + p_c y_j^j \right) \int_{\vec{\eta}} \left( \prod_{j=1}^{M} e^{e_j^j / \sigma_n} \right) \left( \sum_{j=1}^{K} \frac{e^{e_j^j / \sigma_n}}{\sigma_n} \right)^{-M} f(\vec{\eta}) d\vec{\eta}
\]

### F Details of the channel expenditure model

#### F.1 Likelihood function

From Section 3.4, the objective is to compute the integral in equation (17b):

\[
\ell(\vec{\theta} = \vec{\theta}^\ast | e_c^\ast) = |J_{V \rightarrow e_c^\ast}| \phi \left( \frac{h(e_c^\ast)}{\sigma_v} \right) \int_{\xi_n = \log(b) - V^\ast_c}^{\infty} \left( \prod_{j \neq c} \int_{\xi_j = -\infty}^{\xi_j^\ast - \xi_c} \lambda \left( \frac{\xi_j^\ast - \xi_j}{\sigma_j} \right) d\xi_j \right) \lambda \left( \frac{\xi_c^\ast - \xi_c}{\sigma_c} \right) d\xi_c
\]

\[
= |J_{V \rightarrow e_c^\ast}| \phi \left( \frac{h(e_c^\ast)}{\sigma_v} \right) \int_{\xi_n = \log(b) - V^\ast_c}^{\infty} \left( \prod_{j \neq c} \Lambda \left( \frac{\xi_j^\ast - \xi_j + \xi_c^\ast}{\sigma_j} \right) \right) \lambda \left( \frac{\xi_c^\ast}{\sigma_c} \right) d\xi_c
\]

\[
= |J_{V \rightarrow e_c^\ast}| \phi \left( \frac{h(e_c^\ast)}{\sigma_v} \right) \int_{\xi_n = \log(b) - V^\ast_c}^{\infty} \exp \left( - \exp \left( - \frac{\xi_c^\ast - \xi_c}{\sigma_c} \right) \sum_{j \neq c} \exp \left( \frac{\xi_j^\ast - \xi_j}{\sigma_j} \right) \right) d\xi_c
\]

The final step of the likelihood calculation is tedious but may be easily verified using symbolic processors such as Mathematica. Note that because this portion of the likelihood is conditioned upon positive expenditure in some channel, the Jacobian always appears and is computed as:

\[
J_{V \rightarrow e_c} = \frac{\partial \nu}{\partial e_c} = \frac{\partial h(e_c)}{\partial e_c} = (1 + r_c) \left( 1 - \frac{\nu_c(e_c)}{\nu_c(e_c)} \right) \left( f_c + (1 + r_c) \left( e_c^\ast + \frac{1}{\nu_c(e)} \right) \right)^{-1}
\]