The Impact of Advertising along the Conversion Funnel*

Stephan Seiler
Stanford University

Song Yao
Northwestern University

This draft: February 20, 2017

We assemble a unique dataset that combines information on supermarket feature advertising with path-tracking data on consumers' movement within the store as well as purchase information. Using these novel data, we trace out how advertising affects consumer behavior along the path-to-purchase. We find that advertising has no significant effect on the number of consumers visiting the category being advertised. The null effect is precisely estimated, and even at the upper bound of the confidence interval, a one-standard-deviation shift in advertising increases category traffic by only 1.3%. On the other hand, we do find a significant effect at the lower end of the conversion funnel. A one-standard-deviation change in advertising (evaluated at the point estimate) increases category-level sales by 10%. We further decompose the impact on sales and find the increase is driven by consumers buying a larger number of different products within the category. The number of consumers purchasing in the category is not affected. We find no evidence of spillover effects of advertising between categories that are stocked in proximity of each other in the store, and find positive spillovers between products in the same category. Based on these patterns, we propose a simple model that explores potential mechanisms.

Keywords: Advertising, Conversion Funnel, Spillovers, Path-tracking Data

*We thank Eric Anderson, Tomomichi Amano, Eric Bradlow, Pradeep Chintagunta, Yufeng Huang, Carl Mela, Sridhar Narayanan, Aviv Nevo, Thomas Otter, and Brad Shapiro for their feedback. We particularly thank Pedro Gardete for helping us compute wild bootstrap standard errors. All errors are our own.

1 Introduction

Advertising is an important part of firms' marketing mix and plays a crucial role in many industries. In the retail industry, advertising spending was over 10 billion dollars in 2014, which equals 1.4% of gross sales (in comparison, net profits were 1.5% of gross sales).[1] Feature advertising, whereby stores promote specific products using newspaper inserts and store fliers, accounts for 42% of the advertising budget and is the focus of our analysis. Given the large amount of spending, marketers must understand the effectiveness of feature advertising at different stages of the conversion funnel. Managers need to know: Does advertising bring consumers to the store? Does it increase traffic to specific parts of the store? Does advertising increase product sales? Understanding which parts of the conversion funnel are affected is relevant for managers in assessing the overall impact of their advertising efforts. If advertising a specific product brings additional consumers to the store, those additional consumers also likely purchase other (non-advertised) products in the store. If, instead, advertising is most effective at the lower part of the conversion funnel and increases purchases by consumers that would have visited the category even in the absence of the ad, the impact of advertising will be narrower and confined to the category or even only the specific product being advertised.

Despite the practical relevance of decomposing the conversion funnel and the perception that advertising can benefit the store more broadly by increasing traffic (Bodapati and Srinivasan (2006), Chan, Ma, Narasimhan, and Singh (2006)), the empirical evidence on this issue is relatively scant. We posit that this sparsity in research is most likely due to a lack of appropriate data. Marketing researchers have traditionally only been able to observe consumers' purchases, and hence most of the advertising literature focuses on analyzing the impact of advertising on sales, but does not explore the different stages of the conversion process.

In this paper, we make use of a novel data set that allows us to observe consumer behavior in a brick-and-mortar store at a greater level of detail.[2] Specifically, we use a data set of consumer "path-tracking" information obtained from radio-frequency identification (RFID) tags that are attached to consumers' shopping carts. This data set allows us to track precisely which path the consumer took through the store as well as where she was located in the store at each point in time. Combined with data on product locations, this approach allows us to measure whether a consumer visited a particular product category, at what time during her trip she made the visit, and how much time she spent in front of the shelf. For the same set of consumers, we also observe purchases as well as the feature advertising they were exposed to across a large set of categories. Using all these pieces of data together allows us to investigate the impact of advertising onto parts of the consumer's decision process that are typically not observed.

More specifically, the research questions we address in this paper are the following. First, we

[1] "The Food Industry Speaks 2015," Food Marketing Institute.

[2] We note that despite the increasing importance of online retail, brick-and-mortar stores still maintain a 90% market share in the retail market (see https://www.internetretailer.com/trends/sales/e-commerce-share-total-us-retail-sales-2015-2011/, accessed on 12/15/2016).

analyze at what stage of the conversion process advertising has the largest impact on consumers. Second, we analyze the consequences of this decomposition in terms of cannibalization and/or spillover effects from advertising. To answer the first question, we make use of the path-tracking data on consumers' movement within the store and analyze whether advertising affects the number of consumers visiting a particular category. We then analyze the impact of advertising on purchase behavior conditional on visiting the category. To answer the second question, we analyze spillover effects across categories that are stocked near each other in the store, as well as between individual products within each category.

Our paper establishes several key findings. First, we investigate *whether advertising drives foot traffic* to the advertised category. We implement this analysis by regressing the number of consumers visiting a specific category on a given day on the number of advertised products in that category, while controlling for category fixed effects and other marketing activity, namely, price reductions and product displays. Surprisingly, we find that feature advertising *does not* increase traffic toward featured categories. The null effect is precisely estimated and, even at the upper bound of the confidence interval, a one-standard-deviation shift in the number of advertised products increases daily category traffic by only 1.3%. Therefore, any possible increase in sales must be driven by an effect of advertising on purchase behavior conditional on visiting the category.[3]

Second, we analyze *whether (and how) advertising affects purchases*. We find the number of advertised products in the category has a significant impact on category sales. A one-standard-deviation increase in advertising leads to a 10% increase in purchase quantity. When decomposing the effect, we find the increase in sales originates almost entirely from one specific margin of adjustment: We find advertising does not significantly affect the number of consumers purchasing in the category. Instead, the quantity purchased per consumer is the primary driver of the sales increase. Furthermore, the increase in sales is driven by consumers purchasing a larger number of different products from the same category, rather than multiple units of the same product. Together with the results from the traffic data, this sales decomposition paints a detailed picture along which margins advertising is able to affect consumer behavior. We find that advertising does have a significant impact, both in a statistical and economic sense, on the final outcome variable of interest, quantity sold. However, along the conversion funnel, advertising is ineffective at various stages of the process. It does not affect traffic patterns, nor does it convert a higher number of consumers to buying in the category. Instead, the overall advertising effect is primarily driven by consumers purchasing additional products within the category.

Third, we investigate advertising spillover effects both within categories (between different prod-

[3] Our findings regarding category traffic do not directly speak to the effect on store traffic. However, based on our findings, weekly variation in advertising is unlikely to affect store traffic. If advertising for a specific category does drive additional consumers to visit the store, we would expect these additional consumers to visit the advertised category, and hence such a store traffic effect should be detectable as an increase in category foot traffic, which we do not find. However, advertising may affect store traffic in the long-term by affecting the price image of the store (Mela, Gupta, and Lehmann 1997, Jedidi, Mela, and Gupta 1999) or generally by building the brand of the store as a whole. The analysis of such effects that are not specific to the categories being featured is beyond the scope of this paper. We return to the issue of advertising impact on store choice in Section 3.3.

ucts) as well as across categories that are stocked in proximity of each other in the store. We implement the latter analysis based on a detailed map of the store that allows us to define the location as well as the set of nearby products for each category. We find no evidence that advertising in a specific category increases purchases in other nearby product categories. However, within categories, we do find evidence of positive advertising spillovers between individual products. The impact of advertising on other products in the same category is substantial and the total category-level impact of advertising is three times as large as the effect on the advertised product itself. Our results thus imply advertising can be used to promote the category as a whole, rather than just the individual product being featured. At the same time, the impact of feature advertising does not spill over to other nearby categories and is hence confined to the specific category being advertised.

Finally, we use a simple model to tie the various findings together and to explore the behavioral mechanisms that are consistent with the data patterns. We posit two alternative scenarios that can explain our set of findings and especially the fact that advertising is only effective at the lower end of the conversion funnel. One possible explanation is that the consumer might be exposed to an ad without taking any immediate action. Instead, she only retrieves the memory of the ad when she is in front of the aisle and interacts with the advertised category. This type of memory retrieval based on an external stimulus can thus explain the presence of an effect only at the lower end of the conversion funnel. Alternatively, it could be the case that only consumers that were already planning to purchase in the category choose to pay attention to the ads for the specific category. Such self-selection into advertising consumption will similarly lead to an absence of an effect in earlier stages of the conversion funnel.

Our paper contributes to several strands of literature. First, it extends the work using data on consumers' within-store movement, such as Hui, Bradlow, and Fader (2009) who document shoppers' deviations from the most efficient path through the store, and Jain, Misra, and Rudi (2014) who analyze the impact of consumers' interaction with a sales representative on their in-store behavior. Also within this literature, Seiler and Pinna (2016) estimate the benefits from search in terms of price saving from longer in-store search. Hui, Huang, Suher, and Inman (2013) and Hui, Inman, Huang, and Suher (2013) both analyze unplanned shopping behavior using video-tracking and RFID tracking technology, respectively. To the best of our knowledge, none of the prior papers in this literature combined advertising data with data on consumers' movement within the store. Only Jain, Misra, and Rudi (2014) investigate the impact of a marketing intervention, in their case, the interaction with a sales person, onto consumer behavior. As we argue in this paper, understanding how marketing activity affects consumers' path-to-purchase can yield important new insights, and we see this paper as a first foray into this research area.

Apart from path- or video-tracking studies in a brick-and-mortar store context, another application of similar methods is from online browsing data. A range of papers have investigated consumer search behavior in this realm (see, e.g., Kim, Albuquerque, and Bronnenberg (2010), De Los Santos, Hortacsu, and Wildenbeest (2012), Bronnenberg, Kim, and Mela (2016), and Chen and Yao (2016)), but mostly focus on estimating the primitives of the search process such as consumer search costs and preferences. The impact of advertising and other marketing tools is not typically the focus of the analysis. However, the type of analysis we conduct in this paper could also be applied to an online context, which would be an interesting area for future research. We conjecture that certain patterns we find in the physical store setting of our paper might look different in an online context due to the fact that navigating through a brick-and-mortar store is more costly and less flexible than online browsing.[4]

A third stream of literature that we contribute to is the literature on measuring advertising spillovers. Sahni (2016) quantifies spillovers effects in the context of online advertising and analyzes between which types of products spillovers tend to occur. Lewis and Nguyen (2014) provide evidence for spillovers to competing firms in online search behavior following an ad exposure. Anderson and Simester (2013) show spillovers exist for products sold by catalog and that they are most prevalent in categories with higher switching costs. Shapiro (2016) estimates spillovers in the context of pharmaceutical advertising. Sahni, Zou, and Chintagunta (2016) find that email coupons generate spillovers to products to which the coupons do not apply. By contrast, in this paper, we find spillovers to be relatively confined to a specific product category, and they do not spill over to other categories in the store. Similar to Anderson and Simester (2013) and Sahni (2016), we thus characterize the nature of spillovers, in our case, as a function of their proximity to each other in the store. We are furthermore able to analyze the antecedents of spillover effects (or the absence of such effects) by analyzing consumers' movement in the store.

Fourth, our paper relates to studies on the effects of marketing activity on shopping behavior. The set of papers that estimate demand models that include advertising and other marketing-mix variables is large and too extensive to enlist here. A few examples of papers that, like our study, focus on the effect of marketing variables at different levels of aggregation are the following. In the context of price promotions, Kumar and Leone (1988) and Walters (1991) find significant substitution effects between products in the same store, but results regarding across-store substitution are mixed. Walters and MacKenzie (1988) also find significant within-category substitution but only weak support for an increase in total store-level sales due to promotions for individual products.

Finally, we contribute to the literature on measuring advertising effects more broadly (see the summary in Bagwell (2007)). However, we differ from most of the prior literature by focusing not only on the impact of advertising on purchases, but also on consumers' movement through the store.[5] Furthermore, we provide a decomposition of the sales effect of advertising into the impact on the number of consumers purchasing in the category, as well as the number of products and quantity purchased per consumer.

The remainder of the paper is structured as follows. In Section 2, we present the data and descriptive statistics. In Section 3, we analyze the impact of advertising on category traffic and sales. In Section 4, we discuss identification and provide an extensive set of robustness checks.

[4]One paper that does investigate the role of advertising along different stages of the consumer's decision process (in the context of financial services) is Honka, Hortacsu, and Vitorino (2016).

[5]Johnson, Lewis, and Nubbemeyer (2016) investigate the impact of advertising along the conversion funnel in the context of online shopping.

Sections 5 and 6 investigate spillover effects and present additional results. Section 7 discusses the possible underlying mechanisms that are consistent with our empirical findings, using a simple model. Section 8 concludes.

2 Data and Descriptive Statistics

Our data come from two sources. First, we obtained data from a large store in Northern California that belongs to a major supermarket chain.[6] For this store, we observe individual-level purchases as well as data on the path a consumer took through the store for a subset of shopping trips over a period of 26 days.[7] In terms of the purchase data, we have information on all consumers that visited the store during these 26 days. For each shopping trip, we observe the full basket of products as well as the price paid for each item. Furthermore, using the sales data, we are able to link the path data to the corresponding purchase baskets.[8] Finally, we have detailed information on the location at which each UPC is stocked in the store. We complement these data with a second piece of data containing information on feature advertising from the IRI data set (see Bronnenberg, Kruger, and Mela (2008)). Below, we provide more details on the path data as well as how the feature advertising and path data are merged to form the final data set.

2.1 Path Data

We record the paths consumers take when walking through the store, using RFID tags that are attached to their shopping carts and baskets (see Sorensen (2003) and Hui, Fader, and Bradlow (2009)). Each RFID tag emits a signal approximately every four seconds that is received by a set of antennas throughout the store. Based on the signal, triangulation from multiple antennas is used to pinpoint the consumer's precise location. The consumer's location is then assigned to a particular point on a grid of "traffic points," which are overlaid onto the store map and are about four feet apart from each other, thus allowing for a fairly granular tracking of the consumer. For every path, we observe a sequence of consecutive traffic points that the consumer passed on her shopping trip with a time stamp associated with each point.[9] We also note that not all shopping carts and baskets in the store are equipped with RFID tags, and we therefore only observe path data for 7% of all store visits. We use the path data to derive our key outcome variable: the daily number of consumers visiting a particular product category.

To define category visits, we first find the locations of all relevant UPCs for a given category and

[6] We are not able to disclose the identity of the supermarket. The store has a fairly typical format with a trading area of about 45,000 square feet and a product range of 30,000 UPCs.

[7] The days in the path data are 8/24/2006 - 8/29/2006 and 9/7/2006 - 9/26/2006.

[8] In section (A.1) of the appendix, we provide details on how the two pieces of data are combined.

[9] If a consumer moves farther than to an adjacent traffic point between signals, the movement over traffic points in between the signals is interpolated. Because the signal is emitted at a high frequency, little interpolation is necessary for most trips.

the traffic points associated with the set of products belonging to that specific category.[10] For each shopping trip, we consider the consumer to have visited the category if, during her trip, she was located on a certain number of traffic points associated with the category. In our baseline definition of a visit, we require a trip to cross at least three traffic points pertaining to the category.[11] We also compute how far into the trip the consumer walked past a specific category, as well as how much time the consumer spent at a specific category's location. The former is obtained by calculating the time elapsed between the beginning of the trip and the moment at which the consumer is first located on a traffic point associated with the category. The latter records the total time a consumer spends on traffic points associated with the category.

Figure 1 illustrates the definition of these variables for a specific trip (indicated by the dashed line) and category. The figure depicts an illustrative aisle of the supermarket that stocks the focal category at the lower right-hand side of the aisle. A series of traffic points inside the aisle, as well as at the lower end, are considered to be in the vicinity of the category and are used to identify whether the consumer visited the category. In this example, the consumer passed six traffic points associated with the category, and hence her trip qualifies as a visit to the category. To compute the timing of the category visit, we retrieve the time stamp when the consumer is first located on one of the relevant traffic points (in this case, the lowermost traffic point inside the aisle) and calculate the time elapsed since the start of her trip. Finally, dwell-time is measured by the total amount of time spent on traffic points belonging to the specific category (i.e., the six traffic points in the lower part of the aisle).

We aggregate all three variables to the category/day level for our empirical analysis. In the case of category visits, we calculate the total number of consumers visiting the specific category each day.[12] With regards to visit timing and dwell-time, we compute the average value of the respective variable at the category/day level.

Finally, we define a product "pick-up" as our purchase outcome. A pick-up is recorded if a product is observed in the consumer's checkout basket *and* the consumer visited the relevant category. We therefore do not count purchases for which no path data are available. Hence, both traffic and purchase outcomes are based on the same sample of consumers.

2.2 Feature Advertising Data

We supplement the purchase and path data with additional information on feature advertising, which we obtain from the IRI data set. The store-level IRI data contain purchase information, feature advertising at the product/store/week level, as well as information on price and product displays. We only make use of the IRI data in a limited way by complementing our main data set with the relevant feature advertising information, which is missing from the path-tracking data. A

[10] The data provide the linkage between traffic and product locations. Most product locations are associated with two or three traffic points.

[11] Our results are not sensitive to this assumption, and we provide robustness checks using different visit definitions.

[12] Due to the fact that the traffic data only cover 7% of all store visits, we re-scale the daily traffic count by (1/0.07). Product pick-ups are re-scaled in the same fashion.

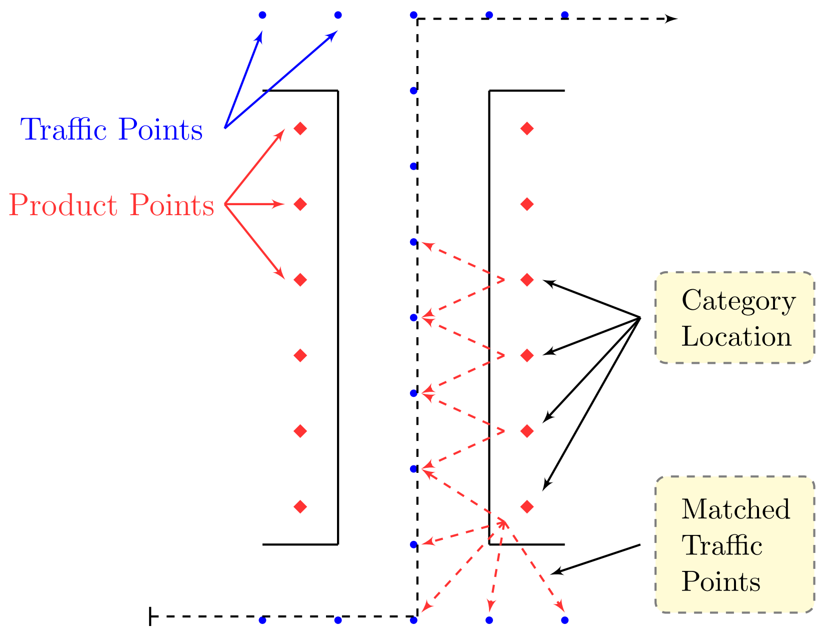

Figure 1: **Category-Visit Definition.** The picture illustrates a consumer traversing an aisle. Consumer location within the aisle is recorded on a grid of traffic points. Products are located at specific locations on the shelf, which are coded up as a grid of product points. Product points are matched to nearby traffic points, allowing us to measure whether and when a consumer visited the category. The dashed black line depicts the consumer's path when traversing the aisle.

product is considered to be featured if it appears in the supermarket's weekly advertising leaflet such as the one displayed in Figure 2 (we display half of a page of a feature advertising leaflet for a store similar to the one in our data).

Unfortunately, the store for which we have the path data is not itself contained in the IRI data, which only contain a sample of stores. However, for the purpose of obtaining information on feature advertising, this issue is not particularly problematic, because stores of the same chain located in the same local market use the identical feature advertising.[13] We are hence able to infer the relevant feature advertising information from several stores of the same chain that are located in the same market (Northern California) and are contained in the IRI data set. We then add the feature advertising information for each UPC/day combination to our path-tracking data set. We note that although most of our analysis is conducted at the daily level, feature advertising only changes at weekly intervals. Our final data set covers four weeks and hence contains four sets of featured products per category. We also note that featured products are not labeled differently in the store (see Figure A1 in the appendix), and hence feature advertising is not correlated with other changes to the store environment that are unobservable to the researcher.

We also use the IRI data to compute a proxy for product displays at the path-data store.

[13]Feature advertising is usually implemented at the market level because doing so allows the chain to only provide one advertising leaflet for the entire market. It is typically paid for by manufacturers and is part of the promotional calendar the retailer and manufacturers agree upon (Mela, Gupta, and Lehmann (1997), Blattberg and Neslin (1990)).

Figure 2: **Example of a Feature Advertising Leaflet (half a page is shown).**

Product displays, in contrast to feature advertising, are often store specific, and hence we cannot perfectly predict product displays from other stores of the same chain. We nevertheless compute a proxy for product displays by calculating the fraction of stores of the same chain that displayed a particular product in a given week, which we interpret as the display probability for the specific product at our focal store. To the extent that stores display similar products,[14] this proxy will allow us to capture the likelihood of a specific product being displayed. We later run a set of robustness checks (in Section 4.3) to assess whether the imperfect measurement of displays affects our results.

[14]The average pairwise correlation of displays (across all categories and weeks) between stores of the same chain in the same market is equal to 0.50.

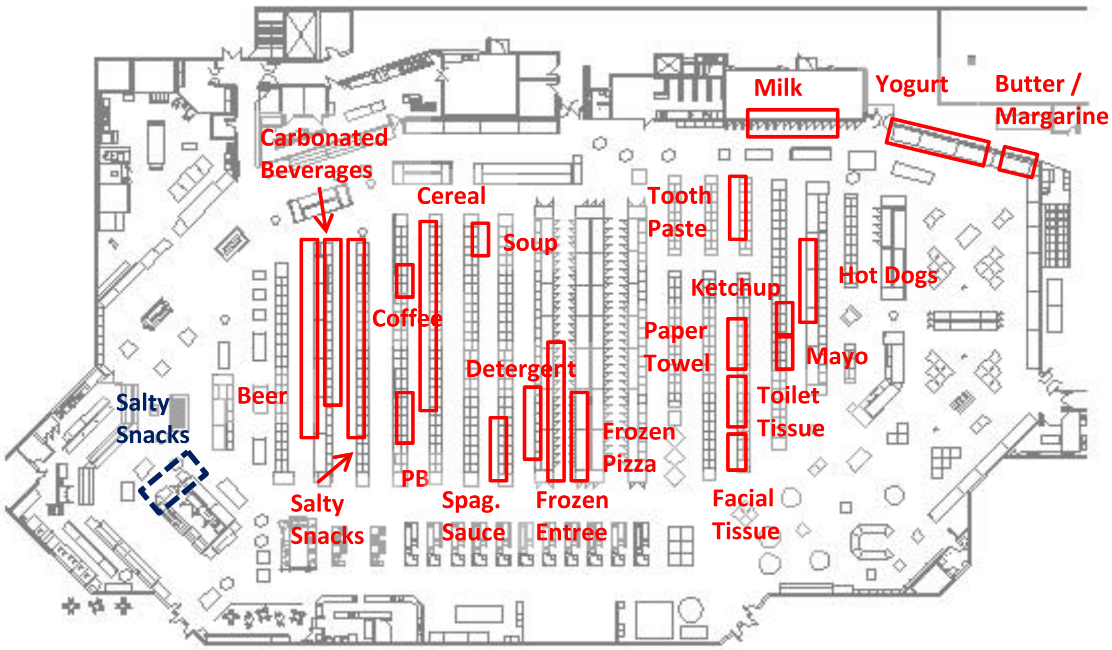

Figure 3: **Store Map with Category Locations.** Primary locations for all 21 categories are depicted using solid rectangles. An illustrative secondary location for the salty snacks category is depicted using a dashed rectangle.

2.3 The Final Data Set: Merging and Category Selection

Our final data set comprises 1,200 products in 21 categories[15] across 26 days. The IRI data set constraints our data set in terms of categories, whereas the path-tracking data limit the time horizon. IRI contains information on 31 categories, but some of those contain very few products that are rarely purchased and other categories are never featured, and thus do not provide relevant variation for our analysis. We therefore end up with 21 categories in our final data.[16]

The primary locations of these categories within the store are displayed in Figure 3. As the figure shows, the categories in our data are fairly spread out in terms of their location, and they also cover a broad set of category "types" such as food and household items, storable and perishable items, etc. The one omission from the set of categories is fresh food, such as produce or fresh meat, which is missing from the set of categories provided in IRI and hence is not part of our analysis. We also note that many categories are stocked in several different parts of the store with a "primary" location in an aisle in the center section of the store as well as additional "secondary" locations in the open areas to the left and right of the primary aisles, as well as on top of the aisles, opposite the entry/exit of the store. The map in Figure 3 depicts all primary locations and an illustrative

[15] The small number of categories prevents us from exploring heterogeneity in the various effects as a function of category characteristics such shelf space, number of brands, and so on.

[16] The categories in IRI not included in our analysis are razors, razor blades, cigarettes, deodorant, diapers, household cleaner, photo, shampoo, sugar substitutes, and tooth brushes.

	Traffic (# Cons. Visiting the Cat.)	Traffic Share (Fraction of Cons. Visiting)	Traffic (Primary Location Only)	Traffic Share (Primary Location Only)	# Cons. Purch. in the Cat.	Total Quantity Purchased	# UPCs	# Feat.
Carb. Bev.	4,866	97.4	1,392	27.8	371	458	108	16.8
Salty Snacks	4,429	88.6	1,237	24.7	182	229	129	18.8
Beer	4,321	86.5	654	13.1	174	187	67	6.0
Soup	3,582	71.7	1,695	33.9	64	153	79	13.3
Spaghetti Sauce	3,534	70.7	1,630	32.6	33	45	43	5.1
Detergent	2,985	59.7	1,052	21.1	17	19	21	3.4
Milk	2,950	59.0	1,062	21.2	90	118	44	4.6
Must./Ketch.	2,616	52.3	685	13.7	17	18	17	0.3
Toothpaste	2,427	48.6	1,022	20.4	13	14	18	0.5
Cereal	2,229	44.6	2,045	40.9	157	238	130	32.0
Frozen Dinner	2,188	43.8	1,305	26.1	103	282	201	66.2
Yogurt	1,988	39.8	1,266	25.3	135	372	143	26.2
Coffee	1,783	35.7	1,783	35.7	21	25	21	1.5
Hot Dog	1,631	32.6	1,548	31.0	18	23	21	2.4
Frozen Pizza	1,359	27.2	1,217	24.4	32	50	44	7.1
Paper Towels	1,123	22.5	1,123	22.5	22	23	16	0.8
Toilet Tissue	1,095	21.9	928	18.6	34	37	15	1.1
Facial Tissue	948	19.0	584	11.7	18	28	13	1.6
Peanut Butter	850	17.0	850	17.0	13	14	17	0.5
Mayonnaise	562	11.2	562	11.2	31	33	20	0.5
Butter/Marg.	165	3.3	165	3.3	18	22	33	1.0

Table 1: **Traffic, Sales, and Feature Advertising Across Categories.** Each column displays the daily average value for the respective variable. The number of UPCs does not vary over time.

secondary location for the salty snacks category in the open area in the left part of the store.

2.4 Descriptive Statistics

We start by providing an overview of the traffic and sales pattern across the categories in our data. The first two columns of Table 1 report total daily category traffic as well as the share of traffic relative to the total number of consumers visiting the store. For simplicity of exposition, the 21 categories are ordered in descending order by their traffic share. We find substantial heterogeneity across categories in terms of the amount of traffic they are exposed to, ranging from over 90% for carbonated beverages to below 10% for butter and margarine. Therefore, consumers rarely visit some parts of the store, whereas other product locations are are passed by a large fraction of consumers.

Columns (1) and (2) are based on all product locations of each category in the store, that is, both the primary location in the aisle as well as any secondary locations. Because secondary locations generally tend to receive more traffic, presumably because they need to be traversed to reach other parts of the store, we also provide an alternative definition of traffic based on primary locations only. Traffic numbers are generally lower, but still vary substantially across the different categories. Also, for some categories, the difference between total traffic and primary location traffic is large, and the gap can be up to 70 percentage points in the case of carbonated beverages. Due to these pronounced differences, we later analyze traffic flows separately for primary and secondary locations. Furthermore, we note that primary locations are typically signposted with the names of the categories stocked in the specific aisle, and hence visits to those locations are likely to be more indicative of consumers explicitly seeking out the category.

We also report category-level sales in terms of total quantity as well as the number of consumers purchasing in the category. We find large heterogeneity in sales levels as well as conversion rates of visiting consumers (captured by the traffic count) to sales.

Finally, we report the number of UPCs in each category and the average number of featured products in the final two columns of Table 1. On average, about 10%-20% of UPCs are featured on any given day within a category. Importantly for our empirical analysis, substantial variation exists in the number of featured products. The standard deviation of the number of featured products in our sample is equal to 17.3, and is 7.9 when isolating only within-category variation in the number of featured products.[17] The average difference between the highest and lowest number of featured products (across the four weeks of our sample) is equal to 12.1 across all categories. We also note that different types of marketing activity are not strongly correlated with each other, and we can hence isolate the effect of feature advertising from the impact of other marketing variables such as promotions and displays (in all of our regressions, we control for other marketing activity). The correlation between the number of feature ads and price promotions (displays) in our sample is equal to 0.35 (0.36) after controlling for category fixed effects. In a larger sample of comparable stores in the IRI data (which we use in a robustness check in Section 4), these correlations are even lower and equal to 0.19 (0.09).

3 Decomposing the Impact of Advertising

3.1 Category Traffic

We start by analyzing the impact of feature advertising on category traffic. As noted earlier, researchers have typically not analyzed this part of the conversion funnel because of the lack of information on consumers' movement within the store. The path-tracking data provide us with a unique opportunity to unpack the effect of advertising by analyzing this "upper level" of the conversion funnel. To the best of our knowledge, this paper is the first to provide such an analysis.

[17] We compute the within-category standard deviation by regressing the number of featured products on category fixed effects. We then compute the standard deviation of the residuals from this regression.

Our empirical strategy is to regress daily category traffic onto the number of featured products within that category, as well as category and day fixed effects, and controls for other marketing activity. Standard errors are clustered at the category level.[18] Formally, we estimate the following regression:

$$Traffic_{ct} = \alpha \times FeatureNum_{ct} + X_{ct}'\beta + \delta_c + \theta_t + \varepsilon_{ct}, \tag{1}$$

where $Traffic_{ct}$ denotes category traffic, that is, the number of consumers visiting category c on day t. $FeatureNum_{ct}$ denotes the number of featured UPCs in category c on day t. X_{ct} denotes a vector of other (time-varying) marketing variables. Specifically, we include the number of promoted items in the category, the average category-level price,[19] and a proxy for the number of displayed items (see Section 2.2). δ_c and θ_t denote category and day fixed effects, respectively. ε_{ct} is the regression error term.[20]

Our baseline specification defines a category visit as a trip that passes at least three traffic points which are associated with the category and is based on all product locations of the category. Column (1) of Table 2 reports the results from this regression. We find that the number of features has *no* statistically significant impact on category traffic (p-value of 0.707).[21] Furthermore, the coefficient on the number of features is not only insignificant, but also small in magnitude. Featuring one additional product leads to 0.631 additional consumers visiting the category. Relative to an average of 2,270 daily category visits, this effect is small.

To further illustrate the magnitude of the effect, consider a one-standard-deviation increase in the number-of-features variable, which is equal to eight additional products being featured.[22] Such an increase in the feature advertising variable leads to about five additional visitors (0.631*8), a 0.22% increase (5/2270). Even evaluated at the upper bound of the 95% confidence interval (i.e., two standard deviations above the point estimate), the effect magnitude is still small. A one-standard-deviation increase in the number of features will lead to 31 additional visitors, a mere 1.3% increase in the number of visits. A final way to assess the relevance of the effect in terms of magnitude is to compare it to the effect of feature advertising onto sales, which we present below. When running the equivalent regression to the one above, but using category sales as the dependent

[18] We also implement the wild bootstrap method that Cameron, Gelbach, and Miller (2008) propose for settings with a small number of clusters. For our baseline regressions (for the impact on traffic as well as sales), we find the level of precision is slightly *higher* when applying the bootstrap procedure.

[19] The price information is obtained from the purchase data. A promotion is defined as a reduction of at least 15% relative to the base price. The average price level is computed as the average (unweighted) price of all products in the category, and captures promotional price fluctuation over time in a more continuous fashion (relative to the number-of-promotions variable).

[20] We would not expect price to be an important control in the traffic regression, because product prices are usually not known to the consumer before reaching the shelf. We nevertheless maintain the number of promotions and average price as control variables in the traffic regression in order to make the regression as comparable as possible with the later regression of sales onto feature advertising and other marketing variables (where price controls are more important).

[21] Using the wild bootstrap procedure (Cameron, Gelbach, and Miller (2008)) we obtain a p-value of 0.645.

[22] We compute the standard deviation of features *within* categories by regressing the feature variable onto category fixed effects and then calculating the standard deviation of the residuals from this regression.

variable, we find a statistically significant increase in sales of 10%. Therefore, the effect of feature advertising on sales (evaluated at the point estimate) is an order of magnitude larger than its effect on traffic (evaluated at the upper bound of the 95% confidence interval).

To probe the robustness of the null result regarding the impact of features onto traffic, we run several additional specifications. We first implement a set of regressions that use different definitions of category traffic. Instead of assuming a consumer visited a category when her path passed at least three traffic points, we consider several more conservative definitions, which require the consumer to pass a larger number of associated traffic points. Columns (2) and (3) of Table 2 report the results from two regressions that base the category-visit definition on at least five and seven traffic points respectively.[23] Results are similar to our baseline specification. The point estimates are close to zero and in fact slightly negative, and the standard errors are smaller relative to the baseline specification reported in column (1). We also run an even larger set of regressions using between 1 and 15 traffic points as the basis for the category traffic definition. Across all 15 specifications, the effect is consistently statistically insignificant with an average (minimum) p-value of 0.841 (0.616) and small in magnitude.

In a second set of robustness checks, we narrow the category definition down to only the primary location of the each category. As described in Section (2.4), many categories are stocked at different locations in the store. Typically, the primary location is either in an aisle or at the back wall of the store (for perishable items), and secondary locations are in the open areas of the stores. Secondary locations often experience higher traffic volume. Furthermore, if consumers who see a feature ad are explicitly seeking out the featured category, we might expect an effect on traffic to show up mostly for the primary locations because those are typically labeled and signposted with the category names. We therefore construct traffic measures using only the primary location of each category for the three, five, and seven traffic-point definitions used previously. The results from those three regressions are reported in columns (4) to (6) of Table 2, and again show a clear null effect with point estimates and standard errors that are of similar magnitude as the estimates in the first three columns.[24]

Furthermore, due to the store layout, some categories may have generally higher traffic than others. As the first column in Table 1 shows, this scenario is particularly true for the traffic definition based on all locations. For instance, over 90% of consumers walk through parts of the store where carbonated beverages are stocked. When products in such a high traffic category are featured, the marginal effect on traffic could potentially be small because of the already high level of baseline traffic. To address this concern, we re-estimate the baseline specification in column (1), but exclude categories with high traffic volume. When we exclude categories with more than 80% or 60% average traffic volume, we again find the effect is small and insignificant.[25]

[23] For categories that have fewer than five or seven traffic points associated with them, a visit was defined as a consumer passing all relevant traffic points.

[24] When basing the category-visit definition on secondary locations only, results are also insignificant.

[25] The coefficient (standard error) is 2.044 (1.745) when using an 80% cutoff and 0.115 (1.157) based on a 60% cutoff.

	All Locations			Only Primary Category Locations		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	# Cat. Visits	# Cat. Visits	# Cat. Visits	# Cat. Visits	# Cat. Visits	# Cat. Visits
Category Visit Definition	$\geq$ 3 Traffic Points Visited	$\geq$ 5 Traffic Points Visited	$\geq$ 7 Traffic Points Visited	$\geq$ 3 Traffic Points Visited	$\geq$ 5 Traffic Points Visited	$\geq$ 7 Traffic Points Visited
Mean	2,270	1,589	1,124	1,133	743	533
S.D.	1,397	1,239	1,100	532	448	425
# Features	0.631	-0.428	-0.295	-0.148	-0.105	-0.250
	(1.654)	(1.029)	(1.005)	(0.963)	(0.969)	(1.059)
Category FEs	Yes	Yes	Yes	Yes	Yes	Yes
Day FEs	Yes	Yes	Yes	Yes	Yes	Yes
Marketing Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	546	546	546	546	546	546
Categories	21	21	21	21	21	21
Days	26	26	26	26	26	26

Table 2: **The Impact of Advertising on Category Traffic.** The unit of observation is a category/day combination. Marketing controls are the number of promoted items in the category, the average category-level price, and a proxy for the number of displayed items. Standard errors are clustered at the category level.

We also note the inclusion of marketing controls (average price level, number of promotions, number of displays) does not play a role in driving the null effect. In Table A1 in the appendix, we present results for a specification without marketing controls as well as specifications that include subsets of those controls. We find our estimate of the effect of feature advertising changes little across those different specifications. This finding is unsurprising because price and displays are only observed once the consumer is in front of the shelf, and we would not expect those marketing covariates to affect traffic.[26]

Finally, we also investigate an alternative measure of category traffic, based on the total number of minutes that consumers spent on traffic points belonging to the category (on a given day), and again find a small and insignificant coefficient estimate.[27]

We conclude that across a wide variety of alternative specifications, the impact of advertising

[26]We note that in the case of the impact on purchases, other marketing activity plays a more important role. Columns (5) to (8) of Table A1 in the appendix report the effect of marketing controls on the impact of feature advertising on purchases. Whereas other marketing activity does have a significant effect on purchases, the coefficient on the number of featured products does not change significantly when further marketing controls are added, which is likely due to the fact that the intensities of different marketing activities are only weakly correlated.

[27]Based on all category locations, the coefficient (standard error) in this specification is equal to 3.722 (4.746). This amount constitutes a small effect relative to the average number of daily category-visit minutes of 1,741.

on category traffic is statistically insignificant and small in magnitude. Consequently, advertising does not seem to be able to attract an economically meaningful number of additional consumers to areas of the store where the specific categories are stocked.

3.2 Category Sales

In this section, we estimate the effect of feature advertising on sales. The objective of this analysis is two-fold. First, we aim to establish whether advertising has any impact on purchases and how the magnitude of the effect compares to the null effect on traffic. Second, we decompose the effect of advertising onto purchases into different adjustment margins. Specifically, we analyze the extent to which the increase in sales originates (1) from a larger number of consumers purchasing in the category, or (2) from individual consumers purchasing a larger quantity of either the same product or of different products from the same category. Together with the results on consumer traffic presented in the previous section, this approach allows us to paint a detailed picture of how advertising affects the consumer's decision-making process.

We start by implementing the following regression:

$$Sales_{ct} = \alpha \times FeatureNum_{ct} + X'_{ct}\beta + \delta_c + \theta_t + \varepsilon_{ct}, \tag{2}$$

where $Sales_{ct}$ denotes a measure of product purchases in category c on day t. $FeatureNum_{ct}$ denotes the number of featured UPCs in category c on day t. X_{ct} denotes a vector of other (time-varying) marketing variables. δ_c and θ_t denote category and day fixed effects, respectively. ε_{ct} is the regression error term. The specification is identical to the one used for analyzing category traffic, but now we use sales instead of traffic as the dependent variable. We note that we "re-use" the same notation for the regression coefficients used in the traffic regression in the interest of simplicity.[28]

To decompose the effect of advertising into the different margins, we use three different measures of purchase outcomes (i.e., $Sales_{ct}$). First, we compute a simple count of the number of consumers purchasing within the category. We then expand this metric to also capture consumers buying different products (UPCs) from the same category as well as consumers purchasing multiple units of the same product. To separately capture both dimensions, we first compute the number of consumer/UPC pairs but ignore multi-unit purchases of the same UPC by the same consumer. In a final step, we also include multi-unit purchases by using total quantity purchased in the category as the dependent variable. To illustrate the decomposition in a simple way, consider a consumer that purchased two units of product A and one unit of product B in the same category. We code this purchase bundle as one consumer, two consumer/UPC pairs, and three units of total quantity purchased in the category.

[28] One could also use the share of purchases divided by the number of consumers visiting the category as the dependent variable. Due to the null effect on traffic, conditioning on category visits will not materially affect the results. For simplicity we therefore focus on the unconditional number of purchases.

Dependent Variable	(1) # Cons. Purchasing	(2) # Cons.-UPC Pairs	(3) Quantity	(4) Quantity From Single-Product Purchases	(5) Quantity From Multi-Product Purchases
Mean	74.4	96.7	113.7	69.5	44.2
S.D.	94.2	123.5	145.2	86.6	88.3
# Features	0.133	1.153**	1.427**	-0.222	1.650***
	(0.194)	(0.469)	(0.599)	(0.210)	(0.553)
Category FEs	Yes	Yes	Yes	Yes	Yes
Day FEs	Yes	Yes	Yes	Yes	Yes
Marketing Controls	Yes	Yes	Yes	Yes	Yes
Observations	546	546	546	546	546
Categories	21	21	21	21	21
Days	26	26	26	26	26

Table 3: **The Impact of Advertising on Purchases.** The unit of observation is a category/day combination. Marketing controls are the number of promoted items in the category, the average category-level price, and a proxy for the number of displayed items. The dependent variable in column (4) (column (5)) is aggregate daily quantity based on all consumers that purchased a single (multiple) UPC(s) in the category. Standard errors are clustered at the category level.

We start by reporting the results from a regression using the count of consumers as the dependent variable in column (1) of Table 3. We find the estimated effect is insignificant, which complements our earlier finding regarding the null effect of features on category traffic. Not only are no additional consumers visiting the category due to advertising, but for those consumers whose shopping paths overlap with the category, advertising also does not convert them into purchasing in the category.

Next, we analyze the impact of advertising on the number of consumer-UPC pairs sold within a category on a given day. The results from this regression are reported in column (2) of Table 3 and show that features significantly increase the number of consumer-UPC pairs. Together with the null result from the first column, this significant result implies advertising leads to the same number of consumers buying a larger number of different products within the category. Finally, we also include multi-unit purchases into our outcome variable by using the total number of purchases as the dependent variable.[29] The results are reported in column (3) of Table 3. The coefficient is statistically significant, and one additional featured product in the category leads to 1.4 additional units being sold.[30] To better assess this magnitude, consider an increase of eight units in the

[29] The data contain outliers for quantity purchased. Some consumers are recorded as having bought up to 80 units of the same UPC. We therefore cap quantity at four units at the consumer-UPC level (the 99th percentile of the quantity distribution).

[30] Using the wild bootstrap procedure, we obtain a p-value of 0.006 for the quantity regression in column (3). Clustered standard errors yield a p-value of 0.027.

number of products featured (a one-standard-deviation shift). Such an increase leads to an 11.4 additional units sold (1.427*8), a 10% increase (1.427*8/114). This effect is large in magnitude and in particular much larger than the corresponding increase in traffic. As reported in the previous section, an additional eight products being featured increases traffic by only 0.22%.

Taken together, columns (1) to (3) show that feature advertising enhances sales by increasing the order size of consumers who are *already* buying, rather than the overall likelihood of purchase across consumers. Furthermore, the increase in order size originates from consumers buying multiple products rather than several units of the same product. We therefore conclude the increased sales come from people buying a larger number of products rather than more units of the same product. To provide further support for the increase in products purchased being the primary driver of the advertising effect, we proceed to decompose the quantity increase documented in column (3) into purchase quantity that originates from single product purchases and quantity purchased that originates from multi-product purchases.[31] Results from these "split regressions" are reported in columns (4) and (5), which show the effect of advertising is driven by an increase in multi-product purchases, whereas the number of single product purchases is unaffected.

3.3 Advertising along the Conversion Funnel

Our findings in the previous two sections show that although advertising leads to an increase in sales, this increase does not originate from more consumers visiting the category or more consumers purchasing in the category. Instead, the effect is driven by individual consumers buying more products from a given category. With regards to the conversion funnel, our findings imply advertising is effective at increasing sales only at the lower end of the conversion funnel; that is, consumers are more likely to purchase an advertised product when they are already in front of the shelf. However, variation in advertising influences no other part of their pre-purchase behavior.

We also note our findings provide some evidence against advertising increasing traffic to the store as a whole, albeit only indirectly. Because we have only one store in our data, it is difficult to directly address the question of whether advertising affects store choice. However, under the assumption that the impact of advertising is weakly increasing throughout the conversion funnel,[32] we can rule out an effect on store traffic due to the absence of a category-traffic effect. In other words, if additional consumers visit the store with the intention to purchase in a specific category as a result of the feature advertising, we would expect these additional consumers to visit the advertised category once they are in the store. The null effect of advertising on category traffic therefore rules out a positive store-traffic effect. However, we cannot rule out that feature ads bring consumers to the store without them necessarily having the intention to purchase a product in the

[31] Column (3) uses total daily quantity aggregated across consumers as the dependent variable: $Q_{jt} = \sum_i q_{ijt}$ where q_{ijt} denotes quantity purchased by consumer i in category j on day t. The dependent variable in column (4) is given by $Q_{jt}^{Single} = \sum_i q_{ijt} * \mathbf{1}(\#UPC_{ijt} = 1)$. Similarly, the dependent variable in column (5) is given by $Q_{jt}^{Multi} = \sum_i q_{ijt} * \mathbf{1}(\#UPC_{ijt} > 1)$.

[32] For example, under this assumption, the effect on category purchases should be larger than the effect on category traffic because it may also enhance conversion among consumers visiting the category.

advertised category.[33]

4 Identification and Robustness Checks

We do not have access to random variation in advertising,[34] and thus the identification of the impact of feature advertising on traffic, sales, and other outcomes relies on variation in marketing activity within categories over time. This empirical strategy leaves two possible factors that could cause bias in our estimates. First, different forms of marketing activity might be correlated over time, and second, advertising could be correlated with time-varying demand shocks (e.g., turkey is more likely to be advertised around Thanksgiving).

A priori, we think both issues are unlikely to be a concern in our setting. First, we control for other marketing activity in all our regressions. Second, our data cover only a short time window and do not contain major holidays or other special events, and hence the scope for demand fluctuations over time is limited. Third, feature advertising is typically determined in advance by the retailer and manufacturers as part of the promotional calendar (e.g., Anderson, Malin, Nakamura, Simester, and Steinsson (2016), Rossi (2014), Quelch and Court (1983)), and is therefore unlikely to be altered in response to short-term demand shocks. Finally, both possible confounds would tend to overstate sales effects, because advertising is most likely positively correlated with demand shocks and other marketing activity. It is, however, less clear how either of the two channels can spuriously generate a null effect on category traffic *and* a positive effect on sales. Nevertheless, we turn to further investigate both issues in a battery of robustness checks below.

We also briefly discuss the related issues of the role of other marketing activity that is delivered at a more aggregate level, such as TV advertising, as well as the impact of measurement error in the variables used in our baseline regressions.

4.1 Time-Varying Demand Shocks

One possible issue in our context could be the fact that within categories, demand varies over time in a way that is correlated with the intensity of feature advertising. To control for such demand shocks in a flexible way, we would ideally want to include category/time-period specific dummies in the regression. However, the unit of observation in our data is a category/day combination, and hence we are not able to control for demand fluctuations at such a granular level.

To circumvent this shortcoming of our main data set, we use additional data for multiple stores from the IRI data (for the same set of categories and over the same time period). The advantage of the IRI data lies in the fact that we have sales and marketing information for a large set of stores

[33] Such a store-level effect can occur if feature advertising affects the general price image of the store (Mela, Gupta, and Lehmann (1997), Jedidi, Mela, and Gupta (1999)), but does not drive consumers to purchase in the specific categories being advertised.

[34] Although randomized field experiments have gained in prominence, they are mostly confined to online markets. To the best of our knowledge, no advertising experiment has been implemented in a brick-and-mortar setting.

(rather than just the one store used in the main regressions).[35] This fact allows us to control for marketing activity, while at the same time being able to back out category-specific time trends that are common across stores. Specifically, we run the following regression:

$$Sales_{sct} = \alpha \times FeatureNum_{sct} + X'_{sct}\beta + \xi_{ct} + \lambda_{sc} + \varepsilon_{sct}, \tag{3}$$

where s denotes a specific store, c denotes the category, and t denotes a week (IRI reports data at the weekly rather than daily level). $Sales_{sct}$, $FeatureNum_{sct}$, and X_{sct} are defined as before, but are store-specific now. X_{sct} contains the number of promoted products, average price, and the number of products on display.[36] Because we have store-category-week-level data, we can allow for category-week-specific demand shocks ξ_{ct}. Furthermore, we also control for store-category fixed effects λ_{sc}. Having recovered the demand shocks from the IRI data, we then include the fitted values $\hat{\xi}_{ct}$ into our baseline regression for the focal store.[37]

We report results with the demand shock as additional control variable for both traffic and sales regressions in columns (3) and (4) of Table 4, which are based on the specifications in column (1) of Table 2 and column (3) of Table 3, respectively. For easier comparison, we also replicate the baseline results for the impact on traffic and sales in columns (1) and (2) of Table 4. The impact of including the demand-shock control on the feature-advertising coefficient in both regressions is minimal, and the null result for traffic as well as the positive and significant effect on sales are robust to the inclusion of this additional variable. Furthermore, the effect of the market-level demand shock on sales is positive but statistically insignificant.[38] Although we only report one traffic and sales-based regression respectively, results are similar when we control for demand shocks using any of the other specifications reported in Tables 2 and 3.

This robustness check deals with the issue of time-varying demand shocks as long as those shocks are common across stores, such as high demand for turkey at thanksgiving. Although we think this assumption is likely to be reasonable, this robustness test is not able to deal with store/category/time-specific demand shocks. The fact that we use IRI data from similar stores in the same geographical market lends further support to the assumption of common demand patterns.

[35] We re-iterate that the main regressions are based on data from only one store, because the path-tracking data were only collected for one store.

[36] We note that in our main data, the display information is noisy because we infer it from other stores of the same chain. The display variable in the IRI data is measured without noise, because the IRI data set contains the exact display information for the stores used in this regression. Importantly, the display information in the IRI data is *not* imputed, but collected at the individual store level.

[37] We use only a subset of comparable stores from IRI. Specifically, we select all stores in California that belong to one of four major chains. Our focal store is also located in the same geographical area and belongs to one of the four chains.

[38] Because the across-store regression is estimated at the weekly level, we divide the demand shock values by 7 in order to make them comparable to the daily sales values used in the baseline regression. We also note the sales regression coefficient on the demand shock is not significantly different from 1, which is the value we would expect to see if the focal store is subject to the same common demand fluctuations as the other stores used to impute the demand shocks.

4.2 Market-level Marketing Activity

Apart from categories and products being promoted at individual stores via feature advertising, several other types of marketing activity for the same products also occur at a higher level of aggregation. Such activity comprises advertising by manufacturers in different media such as TV, radio, and newspaper advertising. Importantly, such advertising is typically delivered at the level of relatively large geographic units (e.g., media markets in the case of TV advertising) and therefore does not vary across stores within a confined geographic area. For this reason, we would expect the category/week demand-shock term estimated from the IRI data in the previous section to also include any demand shifts that such market-level marketing activity induces. Therefore, similar to taste-based shifts of demand over time, any variation in marketing activity over time that is common across stores in the same local market will be controlled for via the imputed demand shock.

Another form of marketing that might occur during our sample period is store ads run by the retailer. Such advertising might change traffic to the whole store and is unlikely to have a differential impact across categories. Our baseline regressions control for such time-varying effects that are common across categories via a set of day fixed effects.

4.3 Correlation in Marketing Activity

A further issue could arise from a correlation of feature advertising with other marketing activity, namely, price promotions and product displays. We note that we control for both price and displays in our main regression. In terms of price controls, we include both the average category price level and the number of promoted items. As a first additional test, we more rigorously control for other marketing activity in a non-linear fashion. Specifically, we include second- and third-order terms for all three marketing control variables. Doing so, we find the coefficient on the number of featured items remains insignificant in the traffic regression, and is significant and positive in the sales regressions. For instance, when regressing total category quantity on feature advertising and controls, we find a coefficient (standard error) of 1.41 (0.64) when including higher-order controls versus 1.43 (0.60) in the corresponding main regression (see column (3) of Table 3).

The most problematic element regarding our attempt to control for other marketing activity is arguably the display variable. As mentioned in Section 2.2, we do not observe display information for the focal store. We therefore approximate product displays by calculating the weekly fraction of stores that display a specific product in stores of the same chain in the same local market (Northern California). The product-specific fraction of displays is then added up across products within a category to yield the number-of-displayed-products proxy variable for each category. This variable is a noisy proxy for displays, and we hence run a set of additional robustness checks.[39]

[39] We emphasize that the display variable in the academic IRI data set is recorded for each store individually. Although industry practice is to sometimes impute display information from other stores, we did confirm with IRI directly that the display information is *not* imputed for the data used here.

First, to assess the possible impact of correlated displays on our estimates, it is useful to consider a few basic descriptive statistics on the usage of displays. Among all three observed marketing activities (i.e., displays, promotions, and feature advertising), displays are the least frequently used. Across all products and stores in the IRI sample used in the previous section, the fraction of product/store/week combinations during which each marketing tool is used is as follows: promotions (40%), feature ads (20%), and displays (9%). Furthermore, the correlation between the different marketing tools is not particularly high. The correlation of feature ads and displays at the category/store level is equal to only 0.09 after controlling for category/store-pair fixed effects.[40] Therefore, the potential for display mis-measurement to bias our estimated effect of feature advertising onto sales and traffic is not particularly large.

Nevertheless, we implement a set of additional regressions to assess possible effects from the imperfect display control variable on our regressions. One thing to note is that the impact of displays onto the traffic and sales regression might be different. In case of the sales regression, one might worry that when controlling imperfectly for displays, the effect of features might be overestimated. In the case of the traffic regression, the direction of the bias is less clear. Conceivably, displays occur in other parts of the stores rather than the typical location of the category. End-of-aisle displays are the most prominent example. Therefore, we might record fewer consumers walking past a specific category because they are able to pick the product up elsewhere. This specific mechanism could therefore lead to a bias toward zero in the traffic regression and an upward bias in the sales regression. We implement two tests below to address this issue.

First, we note that the sales variable used in our estimation is based on product "pickups" (see Section 2.1). In other words, we are only recording the sale of a product if the product appears in the consumer's checkout basket *and* the consumer walked through the aisle where the product is stocked. Therefore, if displays divert consumers away from aisles, because consumers pick up the product elsewhere, both the traffic and sales results will be biased towards zero. Hence, the divergence of traffic and sales effects cannot originate from this mechanism.

We nevertheless further probe our data to test whether displays do lead to more purchases from temporary locations rather than the main category locations. We can implement such a test by computing for each category/day-pair the number of purchased products (from the checkout data) as well as the number of items picked up from the typical category locations. If displays divert traffic and sales away from the typical category locations, we would expect the number of pickups relative to purchases to decrease as a function of displays. We test this hypothesis by regressing the ratio of pickups to purchases onto features, displays, and other control variables (following the specification used in our baseline regressions). We run this regression using the ratio of both primary location and all location pickups relative to total purchases. The results are reported in columns (5) and (6) of Table 4. We find that in both cases, the display proxy variable (not reported in the table) has no significant effect on the pickup/purchase ratio, and feature advertising (which

[40] To control for category/store fixed effects, we regress feature and display dummies respectively onto category/store-pair fixed effects and then analyze the correlation between the residuals from both regressions.

	Baseline		*Demand Shock Control*		*Only Perishable Categories*		*Pickup/Purchase Ratio*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	# Category Visits	Quantity Purchased	# Category Visits	Quantity Purchased	# Category Visits	Quantity Purchased	Pickup to Purchase Ratio (Prim. Loc.)	Pickup to Purchase Ratio (All Loc.)
Category Visit Definition	All Locations $\geq$3 Traffic Points Visited	n/a	All Locations $\geq$3 Traffic Points Visited	n/a	All Locations $\geq$3 Traffic Points Visited	n/a	n/a	n/a
# Features	0.631	1.427**	0.963	1.568***	-0.998	2.085***	0.0003	0.0017**
	(1.654)	(0.599)	(1.569)	(0.472)	(1.261)	(0.301)	(0.0007)	(0.0008)
Imputed Demand Shock			0.872	0.369				
			(0.642)	(0.439)				
Category FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marketing Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	546	546	546	546	156	156	546	546
Categories	21	21	21	21	6	6	21	21
Days	26	26	26	26	26	26	26	26

Table 4: **Robustness Checks.** The unit of observation is a category/day combination. Marketing controls are the number of promoted items in the category, the average category-level price, and a proxy for the number of displayed items. Standard errors are clustered at the category level.

might be correlated with unobserved display variation) has a marginally significant effect in one of the two specification. However, the effect is small in magnitude,[41] and the sign of the effect is *positive*; that is, features led to more pickups relative to total purchases, whereas the display correlation would imply a negative effect. Hence, these regressions provide evidence against displays diverting traffic away from permanent locations. If anything, displays seem to lead to more pickups relative to total purchases.

In a second test, we confine our analysis to perishable products, for which displays are rare due to the need for these products to be stored in refrigerators, and hence they cannot be moved to different locations such as end-of-aisle placements. We re-run the main traffic and sales regressions based on categories containing perishable products only,[42] and find the results are robust to using this sub-sample of categories. Results for both traffic and sales regressions are reported in columns (5) and (6) of Table 4. As for the previous robustness checks, we only report one traffic- and sales-based regression, respectively. However, results are similar for the other specifications reported in Tables 2 and 3.

Finally, one might also worry that the impact of feature advertising on sales is incorrectly attributed to the effect of displays that are located at the category's main location. The robustness test based on perishable products, which are rarely displayed, provides evidence against such a scenario. Furthermore, we find a significant effect on sales of similar magnitude (relative to our baseline regression for the focal store) when estimating our main regression based on the IRI data, which contain a correctly measured display variable.[43]

Taken together, the set of tests described above help rule out bias in our estimates stemming from imprecise measurement of product displays.

4.4 Measurement Error

Finally, we assess the potential for measurement error to bias our estimates. The two variables that are most likely to be subject to measurement error are the category-traffic count and the number-of-displays proxy variable. Importantly, our key explanatory variable, the number of featured products in the category, is less likely to be mis-measured, and hence typical concerns about attenuation bias do not apply here. Instead, any concerns about mis-measurement apply only to a control variable (displays) and the dependent variable in the traffic regressions. As we outline below, measurement error in either case is less problematic.

With regards to (classical) measurement error in the traffic count,[44] some amount of mis-measurement is likely. The primary source of such error occurs due to consumers leaving their

[41] The standard deviation of the pickup/purchase ratio (based on all locations) is 0.306, and hence a one-standard-deviation shift in the number of features (eight additional products) leads to an increase of 4% of a standard deviation (0.0017*8/0.306).

[42] The six perishable categories are frozen entrees, frozen pizza, milk, yogurt, butter / margarine, and hot dogs.

[43] The IRI-based regression referred to here is the one presented in equation 3 in the previous section. The coefficient (standard error) for the effect of feature advertising on sales at the daily level based on the IRI data is equal to 1.077 (0.105).

[44] All of our discussion in this section focuses on classical, that is, additively separable, measurement error.

carts behind while visiting a specific category. However, traffic is used as a dependent variable, and hence any measurement error in traffic will decrease the degree of precision of the regression (by inflating standard errors), but will not lead to biased estimates. As we outlined in detail in Section 3.1, the estimated coefficient on feature advertising in the traffic regression is fairly precise and the effect size is small even at the upper bound of the confidence interval.

A second variable that might be plagued by measurement error is the display variable we discussed extensively in the previous section. Because we include displays only as a control variable, the impact of measurement error on the main coefficient of interest, feature advertising, is indirect. Nevertheless, measurement error in displays can potentially lead to a biased estimate of the impact of feature advertising. If controlling for displays is important to isolate the effect of feature advertising (because feature ads are correlated with displays), the mis-measured display proxy will not be able to control fully for the variation in the *actual* number of displays.

For several reasons, we think such a scenario is unlikely to be problematic for the traffic and sales results presented earlier. First, as documented above, displays tend to be positively correlated with feature advertising, and we would expect them to have a positive effect on traffic and sales.[45] Therefore, not controlling fully for displays will bias the feature coefficient upwards. Mis-measured displays can therefore not account for the null effect on traffic, but they could lead to an overstatement of the impact of feature ads on sales. The latter, however, is unlikely, because the impact of feature ads on sales using the IRI data, where we can control for displays without mis-measurement, is similar to the effect we find in the data for our focal store. Finally, we re-iterate that displays and features are not strongly correlated (as discussed at the beginning of the previous section), and hence the impact of the display control on the feature effect is likely to be minimal. In this regard, we also note the feature coefficient in the traffic and sales regression does not change much when displays are included as a control variable versus when displays are omitted as we show in Table A1 in the appendix.

5 Spillover Effects

Having established and probed the robustness of the effect of advertising along the conversion funnel, we now turn to analyzing the consequences of this decomposition in terms of spillovers to other products and categories. First, we explore whether advertising in the focal category affects sales in categories that are stocked close to a featured category. This analysis makes use of the detailed information on product locations within the store. To the best of our knowledge, such "micro-geographic" spillovers within a store have not previously been explored, because data on store layout and product locations have not typically been available to researchers. Second, we explore whether, within categories, advertising leads to category expansion, brand substitution, or

[45] As discussed above, one could imagine that displays divert traffic away from main category locations, and hence the impact on traffic might be negative. However, our analysis in the previous section provides evidence against such an effect.

	Cross-Category Spillovers			*Within-Category Spillovers*	
	(1)	(2)	(3)	(4)	(5)
Dependent Variable	Quantity	Quantity	Quantity	Quantity	Quantity
Unit of Observations	Category	Category	Category	Product	Product
Definition of Nearby Products	≤15 Feet All Loc.	≤10 Feet All Loc.	≤10 Feet Primary Loc.	n/a	n/a
Mean	1407	676	203	1.99	1.99
S.D.	1125	592	137	6.11	6.11
# Features	-2.800	-1.200	-0.467		
	(1.732)	(1.191)	(0.596)		
Feature Dummy				0.502***	0.420***
				(0.148)	(0.144)
Fraction of Other Products Featured					0.948**
					(0.446)
Category FEs	Yes	Yes	Yes	No	No
Product FEs	No	No	No	Yes	Yes
Day FEs	Yes	Yes	Yes	Yes	Yes
Marketing Controls	Yes	Yes	Yes	Yes	Yes
Observations	546	546	546	31,200	31,200
Products	n/a	n/a	n/a	1,200	1,200
Categories	21	21	21	21	21
Days	26	26	26	26	26

Table 5: **Spillover Effects across and within Categories.** The unit of observation is a category/day combination in columns (1) to (3) and a product/day combination in columns (4) and (5). Marketing controls are the number of promoted items in the category, the average category-level price, and a proxy for the number of displayed items in columns (1) to (4). In column (5), marketing controls also include "fraction of other" versions of each marketing variable. Standard errors are clustered at the category level in columns (1) to (3) and the product level in columns (4) and (5).

positive spillovers between products. We first turn to spillovers between categories that are stocked close to each other.

5.1 Cross-Category Spillovers

Two unique aspects of our data are that we observe both consumers' movement through the store and product locations. In this section, we leverage the second part of the data in order to explore whether feature advertising has any effect beyond the focal category and, specifically, whether feature advertising positively or negatively affects nearby categories. Based on our previous finding that advertising does not affect the number of consumers visiting a specific category, we conjecture that a spillover effect onto nearby categories is not likely to occur.

Our analysis proceeds in a similar fashion as the analysis of sales within the category (see equation 2), except that we substitute sales of nearby products for sales within the category as the dependent variable. Apart from the change in the dependent variable, we employ the same regression framework as earlier and control for category and day fixed effects and a set of marketing controls. To define which products are stocked near the set of 21 categories for which we observe advertising, we first find all locations at which products of a particular category are stocked. Based on these sets of coordinates for each category, we then find all *other* product locations that are within a certain distance of any product point belonging to the category.

Our baseline specification uses all category locations and defines vicinity as a 15-foot radius around each product location. In other words, for, say, the beer category, we find all locations at which beer is stocked and then draw a 15-foot radius around each location and find all possible locations within this radius at which other products might be stocked. We make sure vicinity is only defined in open spaces of the store; that is, we do not consider adjacent aisles within 15 feet as nearby locations, because they are separated by a wall. Having defined nearby locations, we find all products that are stocked at these locations, and compute the total daily sales volume across all products. We hence end up with a count of sales of nearby products at the daily level for all 21 categories.

The results using this baseline definition are reported in column (1) of Table 5. We find a negative but insignificant effect, which is consistent with our prior that cross-category spillover effects are unlikely to occur. Evaluated at the point estimate, the estimated effect corresponds to a 1.6% decrease in sales per featured product, and the 95% confidence interval ranges from a 3.5% decrease to a 0.3% increase. We note that relative to the traffic regressions, our estimates are noisier. Nevertheless, the range of effect magnitudes within the confidence interval are economically relatively small and we can rule out large positive effects.

We also probe the robustness of our results to alternative definitions of nearby categories. We first narrow the radius to 10 feet, and then also employ a definition that is based only on the primary locations of each category rather than all product locations (using a 10-foot radius). Results from both specifications are reported in columns (2) and (3) of Table 5 and show negative and insignificant effects. In further robustness checks, we use every combination of a 5-, 10-, 15-, and 20-foot radius and primary versus all locations to define the vicinity of categories. We find no significant effect in any of those eight regressions. Finally, we also run a set of regressions (not reported in the table) where we distinguish nearby products by their relationship with the focal category. Specifically, we divide nearby products into substitutes, complements, and unrelated products, and run regressions separately for each type. Consistently across all three product types, we find small and insignificant effects.[46]

[46] We manually code whether categories are substitutes, complements or unrelated to each other. For instance, in the vicinity of beer, one substitute category (wine) is stocked as well as several complementary categories (chips, popcorn, etc.). We also note that the majority of nearby products belongs to unrelated categories (88 percent) and

In summary, our results show that, if anything, small negative spillover effects occur, and feature advertising in a given category causes lower sales in nearby categories. The effects are, however, statistically insignificant and economically small relative to the effect on sales within the focal category. We hence conclude that advertising decisions for individual categories can be taken in isolation without a need to coordinate such decisions across categories.

5.2 Within-Category Spillovers

We next proceed to analyze the response to feature advertising at the individual product level. We have already seen that features lead to higher sales at the category level in Table 3. However, the impact of features on purchases of individual products could be larger, because part of the increase in product sales might be due to consumers substituting away from competing products in the same category. Or, inversely, advertising might lead to positive spillover effects onto other products within the category that are not themselves advertised.

To capture both category expansion as well as brand substitution and spillover effects within one unified regression framework, we propose the following linear regression equation:

$$Sales_{jt} = \alpha_1 Feature_{jt} + \alpha_2 \frac{Feature_{-jt}}{N_{c_j} - 1} + Z'_{jt}\beta + \gamma_j + \vartheta_t + e_{jt}, \tag{4}$$

where $Feature_{jt}$ is a dummy variable equal to one if product j is featured on day t. $Feature_{-jt}$ denotes the number of products featured in category c_j that product j belongs to, but excludes product j itself. The variable therefore represents the number of *other* products featured in the same category. Dividing by $(N_{c_j} - 1)$, the number of other products in the category, yields the fraction of other products featured. Z'_{jt} denotes other marketing controls and contains the same variables as previous regressions, but also includes an analogue to the $Feature_{-jt}/(N_{c_j} - 1)$ term for each of the other marketing variables. γ_j and ϑ_t denote product and day fixed effects, and e_{jt} is the error term. Standard errors are clustered at the product level.

To see why the formulation above is useful for analyzing brand substitution, spillover, and category expansion effects, consider the predicted change in sales when product j is featured. At the individual product level, the change in sales is given by

$$E(\triangle Sales_{jt} | \triangle Feature_{jt} = 1) = \alpha_1.$$

We can similarly compute the predicted change for other products in the same category. For any competitor product in the same category, $Feature_{-jt}$ increases by one unit, and hence the predicted change is equal to $\alpha_2/(N_{c_j} - 1)$. Because $(N_{c_j} - 1)$ other products exist, the predicted change aggregated to the category level is given by

$$E(\sum_{k \in c_j} \triangle Sales_{kt} | \triangle Feature_{jt} = 1) = \alpha_1 + (N_{c_j} - 1)\frac{\alpha_2}{(N_{c_j} - 1)} = \alpha_1 + \alpha_2.$$

only a small subset of products are either substitutes or complements of the focal category.

We would expect the product- and category-level effect to be non-negative, and hence $\alpha_1 \geq 0$ and $\alpha_2 \geq -\alpha_1$. In the case of $\alpha_2 = -\alpha_1$, no category expansion effect exists, and the increase in sales for any product is generated entirely from substitution away from other products in the same category. If $\alpha_2 = 0$, no brand substitution takes place, and the entire increase in product-level sales translates into the category growing by the same amount. The case of $\alpha_2 > 0$ represents a positive spillover effect, and hence the category-level increase is larger than the product-level effect.

We present results from the regression outlined above in Table 5. For ease of exposition, in column (4), we first present results from a regression that omits the variable capturing advertising of other products within the same category. Unsurprisingly, we find that feature advertising has a positive effect on product-level sales. In column (5) of Table 5, we report results for the full specification outlined above. We find that both the coefficient on the feature dummy as well as the coefficient on the fraction of other products featured are positive and statistically significant. Furthermore, although less precisely estimated, α_2 is roughly twice as large as the own-advertising effect represented by α_1.[47]

Therefore, advertising for a specific product does not cannibalize sales from other products in the category but instead leads to an increase in sales of other products.[48] Note such spillovers occur without a change in the number of consumers purchasing in the category (as we documented in section 3.2). Hence, advertising does not bring new consumers to the category, some of whom purchase non-advertised products. Instead, consumers who would have purchased in the category, even in the absence of advertising, purchase a larger number of products. Some of these additional products were, however, not advertised. We conjecture such spillovers are related to closeness in product space between advertised products and the non-advertised products that benefit. For instance, a consumer might see an ad for a specific pale ale beer. Later, when purchasing her usual product from the beer category, the consumer remembers the ad and purchases a pale ale beer, but not necessarily the one being advertised. An exhaustive analysis of the determinants of spillover effects is, however, outside the scope of this paper. Therefore, we do not further explore between which pairs of products spillover effects occur.

6 Additional Results

Before proceeding to the analysis of the underlying mechanism, we provide additional results on two further outcomes. First, we investigate the impact of advertising on the timing of category visits, and second, we analyze whether advertising influences the amount of time spent in front of the category. Both outcomes have not typically been observable to researchers in the past, but can

[47]We also assess robustness of these product-level regressions to the inclusion of the category/week demand-shock control used in section 4.1. Results for the coefficients in columns (4) and (5) are almost unaltered. For the specification in column (5), we obtain a coefficient (standard error) on the feature dummy of 0.422 (0.143) and on the "fraction-of-other-featured-products" variable of 1.111 (0.456) when including the imputed demand shock as an additional control variable.

[48]Shapiro (2016) finds a similar pattern in the context of TV advertising for anti-depression drugs. Sahni (2016) also shows that spillovers exist within categories. Specifically, restaurants that serve a similar cuisine (i.e., that belong to the same category) tend to generate spillovers to each other.

be tracked here based on the path data. Both aspects allow us to provide a more complete picture of the impact of advertising and are useful when analyzing the underlying mechanism in section 7. We present a more detailed analysis with regards to both issues in sections A.2 and A.3 of the appendix, and provide only a summary of the key results here.

We first turn to the timing of category visits. One possible impact of feature advertising might be to make the consumer visit a particular category earlier. This type of effect seems likely if the featured product becomes part of the consumer's set of planned purchases on the specific trip and those planned purchases happen earlier on the trip. To systematically explore the timing of category visits, we compute for each shopping trip the point in time at which the consumer is walking past a specific product category for the first time. We then regress the daily average time since the start of the trip during which a specific category was visited on the number of featured products in that particular category (as well as the usual set of control variables). We implement this analysis using time elapsed before visiting the category, as well as the fraction of shopping time elapsed. We find advertising has a small and insignificant effect. This null result is robust to measuring visits to primary versus all locations, as well to measuring visit timing only for consumers who purchased in the category. We hence conclude feature advertising does not influence when consumers visit a specific category.

Second, we analyze the impact of advertising on the time a consumer spends in front of the particular category. One might expect that an advertised product is picked up faster due to the fact that the consumer already knows about it and has an intention to purchase it. Based on the path data, we calculate the duration a consumer spends in front of the shelf (i.e., on traffic points belonging to the specific category) for each purchase. We aggregate this variable to the day/category level and regress the average dwell-time onto the number of features and control variables. We find no significant effect of feature advertising on search time and hence no evidence that consumers pick up items faster or slower after having been exposed to advertising. We note, however, that our test might be under-powered because we use dwell-time across all pickups as our dependent variable. Conceivably, dwell-time changed for the subset of consumers for whom the ad triggered purchase behavior, but detecting this effect is difficult when averaging across all consumers. Second, our measure of dwell-time across the entire category might be quite broad. We therefore also use an alternative measure that only captures the amount of time spent near the specific product that was picked up (rather than the entire category). This regression also yields an insignificant result.

7 Analysis of Mechanism: A Simple Model of Advertising Impact

In this section, we propose a simple model outlining the possible channels through which advertising could conceivably influence category traffic, category-purchase incidence, and individual product purchases. Our analysis captures some elements commonly used in models of advertising impact, such as a distinction between the impact of advertising on the specific product being advertised versus the category to which it belongs. However, we amend the model framework by allowing advertising to either change consumers' purchase intentions directly (e.g., because the consumer puts the advertised product on her shopping list) or by building memory regarding the advertised product. Memory formation is not always included in models of advertising impact, but plays a crucial role in a subset of papers that borrow insights from cognitive psychology.[49] In our context, memory provides a natural explanation for the differential impact of advertising along the conversion funnel we documented earlier, and hence we opt to distinguish between memory formation and explicit purchase intention. The distinguishing feature between the two is the fact that the consumer can retrieve memory only through some external stimulus that reminds her of the ad. For instance, seeing a specific product that was advertised on the shelf can make the consumer remember an ad to which she was exposed earlier.

Importantly, when thinking about the impact of advertising exposure in the case of feature advertising, we need to consider the fact that reading a feature-advertising leaflet and paying attention to a specific ad is the consumer's decision, and hence certain types of consumers will select into reading specific feature ads. We capture this idea by allowing the exposure of consumer i to an ad for product j to be determined by the following expression:

$$AdExp_{ij} = f(CatIntent_{ic,t-1}, ProductIntent_{ij,t-1}).$$

$AdExp_{ij}$ denotes whether the consumer saw the ad for product j. $CatIntent_{ic,t-1}$ denotes the consumer's intention to purchase in the specific category c to which product j belongs, which arises from consumption needs. $ProductIntent_{ij,t-1}$ denotes the purchase intention for a specific product. Note that we use a $(t-1)$ subscript on the two purchase-intention variables to denote that these variables refer to the consumer pre-ad exposure, and purchase intention might change due to the advertising.

To capture the impact of advertising on consumer choices such as category visits and purchases, we model a set of intermediate measures that are influenced by advertising. First, advertising might influence the consumer's purchase intention at the category or product level. That is, after seeing the ad, the consumer might make a mental note to purchase from the specific category or the specific product that was advertised. We denote purchase intent after ad exposure by $CatIntent_{ic}$ and $ProductIntent_{ij}$ (note the $(t-1)$ subscript is dropped here). Furthermore, we consider the possibility that the ad registers with the consumer without her taking an immediate and conscious action. We denote such a memory effect at the category- and product-level, respectively, as $CatMemory_{ic}$ and $ProductMemory_{ij}$. We note the category- and product-level constructs are not independent of each other, and the consumer cannot intend to purchase a specific product without also intending to buy in the category. For simplicity of exposition and because of their relationship to the category- and product-level regressions presented earlier, we maintain the two separate terms.

[49]Sahni (2015) structurally estimates a model of advertising impact on memory formation. See also the discussion of the role of memory in Bagwell (2007).

We then relate these four measures to the three stages of the conversion funnel we observe: category visit ($Visit_{ic}$), category purchase incidence ($Purchase_{ic}$), and product choice ($Purchase_{ij}$):

$$\begin{aligned} Visit_{ic} &= g_{Visit}(CatIntent_{ic}, CatMemory_{ic}), && (5)\\ Purchase_{ic}|Visit_{ic} &= g_{CatPurchase}(CatIntent_{ic}, CatMemory_{ic}), && (6)\\ Purchase_{ij}|Purchase_{ic} &= g_{ProdPurchase}(ProductIntent_{ij}, ProductMemory_{ij}), && (7) \end{aligned}$$

where product-level constructs are excluded from the category-level outcome equations in the first two lines, and category-level variables do not enter the final equation, which conditions on category-level incidence.

We can now relate the expressions above to the effect (or absence of an effect) of advertising on the various outcomes we documented earlier in the paper. First, we find that neither the number of consumers visiting the category nor the number of consumers purchasing in the category changes as a function of advertising. For the case of category visits given by equation 5, it therefore holds that

$$\begin{aligned} &\frac{\partial Visit_{ic}}{\partial CatIntent_{ic}} \frac{\partial CatIntent_{ic}}{\partial AdExp_{ij}} \\ &+ \frac{\partial Visit_{ic}}{\partial CatMemory_{ic}} \frac{\partial CatMemory_{ic}}{\partial AdExp_{ij}} = 0. \end{aligned}$$

It is natural to assume $\partial Visit_{ic}/\partial CatIntent_{ic} > 0$ because consumers needing to buy a specific category will be more likely to visit it. And hence it follows that advertising exposure does not affect purchase intent at the category level, that is, $\partial CatIntent_{ic}/\partial AdExp_{ij} = 0$.[50] The case is less clear for the memory effect. The derivative of memory with respect to advertising could be zero or, alternatively, memory might be built, but not retrieved, when the consumer enters the store and decides which categories to visit.

A similar reasoning applies to the case of category purchase incidence (equation 6). Category purchase intent likely increases purchase incidence in the category ($\partial Purchase_{ic}|Visit_{ic}/\partial CatIntent_{ic} > 0$), and hence the absence of an effect of advertising on category-level purchase incidence in our empirical analysis provides another piece of evidence that advertising does not alter category purchase intentions. Again, the case for the memory effect is less clear. The null effect is consistent with either no impact of advertising on memory or no memory retrieval when the consumer walks past the category.

Finally, we find that purchases of the specific product being advertised increase (equation 7):

[50] Any of the four terms in the equation is unlikely to be negative, and hence both terms of the sum need to be equal to zero.

$$\frac{\partial Purchase_{ij}|Purchase_{ic}}{\partial ProductIntent_{ij}}\frac{\partial ProductIntent_{ij}}{\partial AdExp_{ij}} + \frac{\partial Purchase_{ij}|Purchase_{ic}}{\partial ProductMemory_{ij}}\frac{\partial ProductMemory_{ij}}{\partial AdExp_{ij}} > 0.$$

In words, the effect on the advertised product could originate from purchase intent regarding the product, holding constant category purchase intent. That is, the consumer already wanted to purchase in the category, and the ad led her to switch to or add a product she would not have purchased in the absence of the ad. Alternatively, a memory effect at the product level could explain the effect. That is, when engaging with the category, the consumer is reminded of seeing an ad, which leads her to purchase the advertised product.[51]

Based on these derivations, two explanations are possible as to why we observe an impact of advertising only at the lower end of the conversion funnel. One possible scenario is that consumers who are exposed to the ad do not take any explicit action and only retrieve the memory of the ad when they are in front of the shelf and engage with the category. This scenario suggests memory retrieval only occurs when the consumer directly interacts with the category, and simply walking past the category is not sufficient to trigger the memory. Such a mechanism is consistent with the literature on memory and retrieval cues (e.g., Keller (1987), Lee (2002), and Lee and Labroo (2004)).

Alternatively, the feature ad might only alter consumers' purchase intentions at the product but not the category level. In other words, consumers who did not intend to purchase in the category will not be converted by the ad, but consumers already wanting to purchase in the category might change their purchase intention with regards to an advertised product. Such an effect could occur if advertising is informative in nature and the consumer does not need to be reminded of the existence of categories, but advertising for an individual product can serve as a reminder for the specific product. An alternative explanation for the absence of a traffic effect is a specific type of selection into ad consumption. Consumers might only pay attention to feature advertising for a category they are already planning to purchase in, and hence $CatIntent_{ic}$ does not increase relative to its pre-exposure state $CatIntent_{ic,t-1}$. However, conditional on purchase intent in the category, advertising might make the purchase of the advertised products more likely. Hence, $ProductIntent_{ij}$ increases relative to $ProductIntent_{ij,t-1}$, and the consumer might purchase additional products within the category on top of the purchase she would have made in the absence of advertising.

The two results regarding purchase timing and dwell-time presented in section 6 provide us with some evidence to distinguish between both mechanisms. If the ad changed the consumer's purchase intent regarding the product, two things are likely to happen. First, the consumer will visit the relevant category earlier on her trip because the advertised product is part of the planned part of her shopping trip. Second, if the consumer already has the intention to purchase the specific

[51] For simplicity, we do not discuss the role of spillover effects in this section and instead focus on the effect for a given product along the conversion funnel.

product, dwell-time in front of the category will be reduced because the consumer does not need to engage in further deliberation in front of the shelf. However, we find no effect on either of those two dimensions, which we interpret as evidence against an effect of advertising on purchase intent. However, we believe that neither test provides conclusive evidence in favor of either mechanism, and we hence remain cautious of ruling out either.

8 Conclusion

In this paper, we leveraged a new data set that combines advertising information with path-tracking data of consumers' movements in a brick-and-mortar store. This unique data set provides a closer look at the different stages of the conversion process that have typically been unobserved.

We find that although advertising has a significant impact on total quantity sold, it is ineffective at various stages of the process. Specifically, advertising does not influence traffic patterns, nor does it convert a higher number of consumers to buy in the category. The null result regarding traffic is precisely estimated, and even at the upper bound of the confidence interval, advertising shows a limited effect on category traffic. Instead, the overall advertising effect is mostly driven by consumers already visiting the category purchasing a larger number of different products within the category. We further investigate the spillover effects of advertising across and within categories. We find that advertising for any specific category has little effect on sales in other nearby categories. Within the same category, we find that a focal product's advertising boosts the sales of other products. Both dimensions of the analysis present an internally consistent picture of advertising impact along the conversion funnel. Advertising does not increase category traffic, and hence the impact of advertising is confined to the category being advertised. At the category level, however, we find advertising does lead to an increase in sales for the advertised product, and a positive spillover effect occurs on other products within the category.

In summary, our study provides a nuanced picture of the impact of advertising along the conversion funnel. Our findings suggest managers need not pay attention to coordinate advertising across categories. At the same time, advertising can benefit the entire category rather than only the advertised product, and hence store managers should think of advertising allocation at the category level, rather than the broader store or narrower product level.

References

Anderson, E., B. A. Malin, E. Nakamura, D. Simester, and J. Steinsson (2016): "Informational Rigidities and the Stickiness of Temporary Sales," *Journal of Monetary Economics*, forthcoming.

Anderson, E. T., and D. Simester (2013): "Advertising in a Competitive Market: The Role of Product Standards, Customer Learning, and Switching Costs," *Journal of Marketing Research*, 50(4), 489–504.

Bagwell, K. (2007): "The Economic Analysis of Advertising," in *Handbook of Industrial Organization*, ed. by M. Armstrong, and R. Porter, vol. 3, chap. 28, pp. 1701–1844. Elsevier Science.

Blattberg, R. C., and S. A. Neslin (1990): *Sales Promotion: Concepts, Methods and Strategies*. Prentice Hall, 2nd edn.

Bodapati, A. V., and V. Srinivasan (2006): "The Impact of Feature Advertising on Customer Store Choice," *Working Paper*.

Bronnenberg, B. J., J. B. Kim, and C. F. Mela (2016): "Zooming In on Choice: How Do Consumers Search for Cameras Online?," *Marketing Science*, 35(5), 693–712.

Bronnenberg, B. J., M. W. Kruger, and C. F. Mela (2008): "The IRI Marketing Data Set," *Marketing Science*, 27(4), 745–748.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 3(90), 414–427.

Chan, T., Y. Ma, C. Narasimhan, and V. Singh (2006): "An Empirical Analysis of Store Competition," *Working Paper*.

Chen, Y., and S. Yao (2016): "Sequential Search with Refinement: Model and Application with Click-stream Data," *Management Science*, forthcoming.

De Los Santos, B. I., A. Hortacsu, and M. Wildenbeest (2012): "Testing Models of Consumer Search using Data on Web Browsing and Purchasing Behavior," *American Economic Review*, 102(6), 2955–2980.

Honka, E., A. Hortacsu, and M. A. Vitorino (2016): "Advertising, Consumer Awareness, and Choice: Evidence from the U.S. Banking Industry," *RAND Journal of Economics*, forthcoming.

Hui, S. K., E. T. Bradlow, and P. S. Fader (2009): "The Traveling Salesman Goes Shopping: The Systematic Deviations of Grocery Paths from TSP Optimality," *Marketing Science*, 28(3), 566–572.

Hui, S. K., P. S. Fader, and E. T. Bradlow (2009): "Path Data in Marketing: An Integrative Framework and Prospectus for Model Building," *Marketing Science*, 28(2), 320–335.

Hui, S. K., Y. Huang, J. Suher, and J. Inman (2013): "Deconstructing the First Moment of Truth: Understanding Unplanned Consideration and Purchase Conversion Using In-Store Video Tracking," *Journal of Marketing Research*, 50(4), 445–462.

Hui, S. K., J. Inman, Y. Huang, and J. Suher (2013): “Estimating the Effect of Travel Distance on Unplanned Spending: Applications to Mobile Promotion Strategies,” *Journal of Marketing*, 77(2), 1–16.

Jain, A., S. Misra, and N. Rudi (2014): “On Search, Sales Assistance and Purchase Decisions: An Analysis Using Retail Video Data,” *Working Paper*, unpublished manuscript.

Jedidi, K., C. F. Mela, and S. Gupta (1999): “Managing Advertising and Promotion for Long-Run Profitability,” *Marketing Science*, 18(1), 1–22.

Johnson, G. A., R. Lewis, and E. Nubbemeyer (2016): “The Online Display Ad Effectiveness Funnel & Carry-Over: A Meta-study of Ghost Ad Experiments,” *Working Paper*.

Keller, K. L. (1987): “Memory Factors in Advertising: The Effect of Advertising Retrieval Cues on Brand Evaluations,” *Journal of Consumer Research*, 14(3), 316–333.

Kim, J. B., P. Albuquerque, and B. J. Bronnenberg (2010): “Online Demand under Limited Consumer Search,” *Marketing Science*, 29(6), 1001–1023.

Kumar, V., and R. P. Leone (1988): “Measuring the Effect of Retail Store Promotions on Brand and Store Substitution,” *Journal of Marketing Research*, 25(2), 178–185.

Lee, A. Y. (2002): “Effects of Implicit Memory on Memory-Based versus Stimulus-Based Brand Choice,” *Journal of Marketing Research*, 39(4), 440–454.

Lee, A. Y., and A. A. Labroo (2004): “The Effect of Conceptual and Perceptual Fluency on Brand Evaluation,” *Journal of Marketing Research*, 41(2), 151–165.

Lewis, R. A., and D. Nguyen (2014): “A Samsung Ad for the iPad? Display Advertising’s Competitive Spillovers to Search,” *Working Paper*, unpublished manuscript.

Mela, C. F., S. Gupta, and D. R. Lehmann (1997): “The Long-Term Impact of Promotion and Advertising on Consumer Brand Choice,” *Journal of Marketing Research*, 34(2), 248–261.

Quelch, J. A., and A. M. Court (1983): “PROCTER & GAMBLE CO. (B),” *HBS Case 584048*.

Rossi, P. E. (2014): “Invited Paper – Even the Rich Can Make Themselves Poor: A Critical Examination of IV Methods in Marketing Applications,” *Marketing Science*, 33(5), 655–672.

Sahni, N. S. (2015): “Effect of temporal spacing between advertising exposures: Evidence from online field experiments,” *Quantitative Marketing and Economics*, 13(3), 203–247.

——— (2016): “Advertising Spillovers: Evidence from Online Field Experiments and Implications for Returns on Advertising,” *Journal of Marketing Research*, 53(4), 459–478.

Sahni, N. S., D. Zou, and P. K. Chintagunta (2016): “Do Targeted Discount Offers Serve as Advertising? Evidence from 70 Field Experiments,” *Management Science*, forthcoming.

Seiler, S., and F. Pinna (2016): “Estimating Search Benefits from Path-tracking Data: Measurement and Determinants,” *Marketing Science*, forthcoming.

Shapiro, B. T. (2016): “Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants,” *Journal of Political Economy*, forthcoming.

Sorensen, H. (2003): "The Science of Shopping," *Marketing Research*, 15(3), 31–35.

Walters, R. G. (1991): "Assessing the Impact of Retail Price Promotions on Product Substitution, Complementary Purchase, and Interstore Sales Displacement," *Journal of Marketing*, 55(2), 17–28.

Walters, R. G., and S. B. MacKenzie (1988): "A Structural Equations Analysis of the Impact of Price Promotions on Store Performance," *Journal of Marketing Research*, 25(1), 51–63.

A Appendix

A.1 Linking Sales and Path Data

One of the important features of our data set is the linkage of sales to trip records. As part of the RFID tracking process, the data report when the consumer arrives at the checkout. Independently, the sales data also have a time stamp for each shopper's transaction at the checkout. Comparing the time stamp of a particular path with the sales data allows us to define a set of "candidate" checkout product baskets that occurred at a similar point in time.[52] Matching which trip goes with which specific transaction involves considering the physical location (i.e., longitude = x and latitude = y relative to the store map) of all the UPCs in each candidate basket. Based on how many of those locations lie on the path we are trying to match, a score is created for the baskets and the highest-scoring one is matched to the path.[53] The matches do not necessarily yield a perfect score, because consumers might occasionally leave the cart and pick up an item. Therefore, we might not see the path of the consumer going past a specific item, even if the item was in her matched purchase basket.

A.2 The Impact of Feature Advertising on Visit Timing

In this section, we describe in more detail the analysis of category-visit timing summarized briefly in section 6. To analyze the timing of visits, we compute for each shopping trip the point in time at which the consumer is for the first time walking past a specific product category. We then compute the average time since the start of the trip during which a specific category was visited at the day/category level.[54] We first regress the time of the visit (measured in minutes since the start of the trip) and fraction of total shopping time elapsed on the number of featured products in a particular category. Both regressions include category and day fixed effects and marketing controls, and hence mirror the traffic regression in equation 1.

We start by implementing the analysis based on all product locations for each category. In other words, we define visit timing as the point in time at which a consumer first passes any location in the store associated with the particular category. The results using both minutes elapsed and the fraction of shopping time elapsed are reported in columns (1) and (2) of Table A2. Columns (3) and (4) replicate the same regressions, but base the visit timing only on the primary

[52] The path-data time stamp that records the arrival at the checkout can be noisy because the consumer will be stationary when standing in line at the cashier. Therefore, checkout baskets within a certain time window after the consumer became stationary in the checkout area qualify as possible matches.

[53] The data provider did not disclose the precise algorithm to us.

[54] We can only define visit timing for consumers who actually pass the category at all during their trip. The day/category average therefore represents the average visit time for the subset of consumers who visit the specific category.

locations of each category. Across all four specifications, we find effects of feature advertising that are consistently small in magnitude and mostly insignificant. Take, for example, the results in column (1). According to the (insignificant) point estimates, a one-standard-deviation increase in the number of features (eight additional features) in a particular category delays the visit to the category by 0.016 minutes (i.e., about 1 second) or shifts the visit timing back by 0.05 percentage points relative to the total time spent in the store.[55] The marginally significant effect in column (4) is similarly small in magnitude and does not constitute an economically meaningful shift in the timing of the category visit.

Finally, advertising might only affect a small set of consumers who are planning to purchase within the category due to the feature ad. When analyzing the visit timing of all consumers in the store, the unaltered behavior of the majority of visitors to the store might mask a significant effect of this group of consumers. We hence isolate the group of consumers who are most likely to be affected, by computing the daily average time of a category visit based only on consumers who purchase in the specific category. The results from regressions based on this measure of visit timing are reported in columns (5) and (6) of Table A2. We again find a null effect of feature advertising on visit timing, and the confidence intervals do not contain effect sizes that are economically important.[56]

We hence conclude feature advertising does not influence when consumers visit a specific category.

A.3 The Impact of Feature Advertising on Dwell-Time

In this section, we provide further details on the impact of advertising on dwell-time in front of the category. Based on the path data, we calculate the total time a consumer spends on traffic points belonging to the specific category for each category in which she purchased during a given shopping trip. Similar to other parts of our analysis, we aggregate this variable to the day/category level and regress the average daily dwell-time onto the number of features (and control variables). Results from this regression are reported in column (7) of Table A2 and show a small and insignificant effect. We note that dwell-time is measured in seconds, and average daily dwell-time has a mean (standard deviation) of 53 (41) seconds. A one-standard-deviation shift in the number of features changes dwell-time by only 0.29 seconds (0.29 = 0.036 * 8).

We note that we would ideally like to measure the time a consumer spent contemplating which product to buy in the category. Total time spent in the vicinity of a given category is likely to be a noisy measure of search time (see Seiler and Pinna (2016) for a detailed discussion of the measurement error associated with path-tracking-based dwell-time measures). We therefore assess robustness of the null effect to using an alternative measure that only captures the amount of time spent near the specific product that was picked up (rather than the entire category). Results from

[55] We also ran the same set of regressions based on distance walked before reaching a specific category (rather than time elapsed), and found similarly small and insignificant results.

[56] The confidence interval for columns (5) and (6), respectively, are equal to [-0.050,0.021] minutes and [-0.089,0.149] percentage points.

this regression are reported in column (8) of Table A2 and also yield an insignificant result and an effect size that is small in magnitude.

Finally, we note that our test might be under-powered because we use dwell-time across all pickups as our dependent variable. Conceivably, dwell-time might have changed for the subset of consumers for whom the ad triggered purchase behavior, but detecting this effect is difficult when averaging across all consumers.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Traffic / *Baseline*				*Sales* / *Baseline*
Dependent Variable	# Category Visits (3 Traffic Point Def.)	# Category Visits (3 Traffic Point Def.)	# Category Visits (3 Traffic Point Def.)	# Category Visits (3 Traffic Point Def.)	Quantity Purchased	Quantity Purchased	Quantity Purchased	Quantity Purchased
# Features	0.431	0.296	0.853	0.631	2.050***	1.501**	1.633**	1.427**
	(1.216)	(1.460)	(1.467)	(1.654)	(0.622)	(0.613)	(0.667)	(0.599)
# Promotions		0.475		1.514		1.518***		1.290***
		(2.212)		(1.504)		(0.449)		(0.301)
Av. Price		9.174		12.724		1.351		0.574
		(60.674)		(59.145)		(9.708)		(9.182)
# Displays (Proxy)			-7.694	-12.757			7.587*	2.794
			(23.503)	(24.551)			(3.889)	(3.899)
Category FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	546	546	546	546	546	546	546	546
Categories	21	21	21	21	21	21	21	21
Days	26	26	26	26	26	26	26	26

Table A1: **Traffic and Sales Regressions: Impact of Marketing Controls.** Columns (4) and (8) represent the baseline traffic and sales regressions. Columns (1) to (4) and (5) to (8) show how results change when using different sets of marketing controls in the traffic and sales regression, respectively. The unit of observation is a category/day combination. Marketing controls are the number of promoted items in the category, the average category-level price, and a proxy for the number of displayed items. Standard errors are clustered at the category level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Minutes Elapsed Since Start of the Trip	Fraction of Time Elapsed (0 to 100)	Minutes Elapsed Since Start of the Trip	Fraction of Time Elapsed (0 to 100)	Minutes Elapsed Since Start of the Trip	Fraction of Time Elapsed (0 to 100)	Dwell-Time (Sec.)	Dwell-Time (Sec.)
Variable Definition	All Locations	All Locations	Primary Location	Primary Location	Conditional on Purchase	Conditional on Purchase	Category Level	Product Level
# Features	0.002	0.006	0.005	0.010*	-0.014	0.030	0.036	-0.015
	(0.002)	(0.005)	(0.003)	(0.006)	(0.017)	(0.057)	(0.133)	(0.016)
Category FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marketing Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	546	546	546	546	546	546	503	503
Categories	21	21	21	21	21	21	21	21
Days	26	26	26	26	26	26	26	26

Table A2: **The Impact of Advertising on Visit Timing.** The unit of observation is a category/day combination. Marketing controls are the number of promoted items in the category, the average category-level price, and a proxy for the number of displayed items. Standard errors are clustered at the category level. Dwell-time is not observed for some day/category combinations. Therefore, the number of observations is smaller in columns (7) and (8).

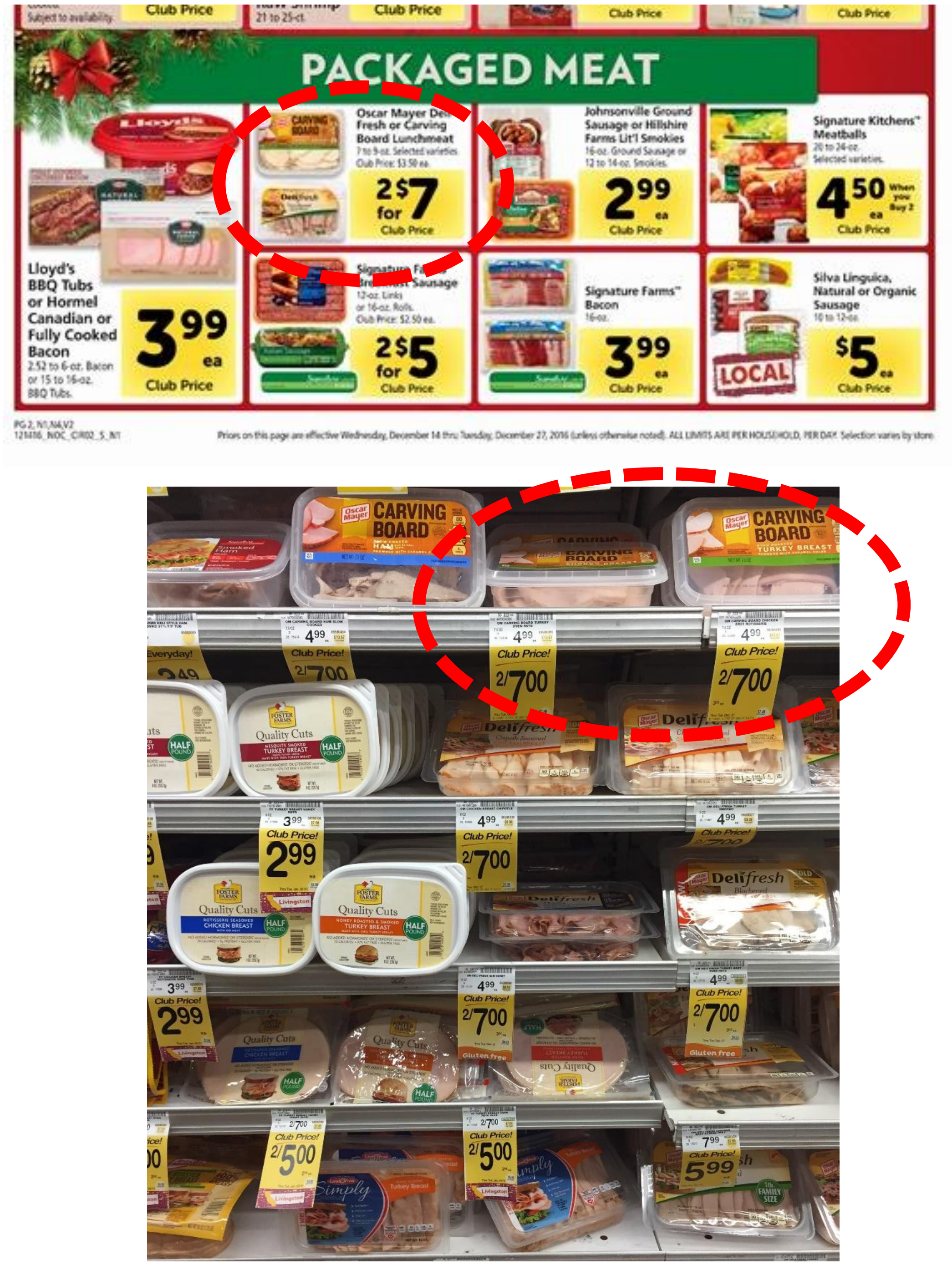

Figure A1: **Example: Feature Advertising and Shelf Labeling.** The top picture shows part of the weekly feature advertising leaflet of a store comparable to the one in our data. The bottom picture shows the labeling on the shelf in the same week. The dashed circle highlights the advertised product. The advertised product is not labeled more saliently. For example, right below the advertised product, another product (which is not featured in the advertising leaflet) has an almost identical label (showing "2 for 7 dollars").