Dynamic Effects of Price Promotions: Field Evidence, Promotion Search and Supply-Side Implications*

Andrés Elberg[†] Pedro M. Gardete[‡] Rosario Macera[§] Carlos Noton[¶]

May 2017

Abstract

We use a large-scale field experiment to investigate the dynamic effects of price promotions. We vary the promotion depth of 170 products across 17 categories in 10 supermarket stores of a major Chilean retailer by randomly assigning at the store level deep and shallow promotions. We find that treated customers, who face deeper promotions in the past, were 12.6 percent more likely than their control counterparts to buy promoted items in a subsequent shallow promotion common to all stores. We rationalize this finding by a sequential search model in which consumers rationally expect better deals after deep discounts given the retailer's historical promotions. The model implies that small manufacturers may benefit from offering unilateral discounts, but may become worse off when large manufacturers are also willing to offer promotions. In cases where all competitors prefer to offer promotions with positive probability, consumer search may make all firms worse off.

*We thank Claudio Palominos and Breno Vieira for outstanding research assistance, as well as the participants in the UCSD Rady Field Experimentation Conference, the Workshop in Consumer Analytics and the SICS Conference, and Bryan Bollinger, Harikesh Nair and Duncan Simester for helpful comments. All errors are our own.

†Universidad Diego Portales, Economics Department; andres.elberg@udp.cl

‡Stanford University, Graduate School of Business; gardete@stanford.edu

§Universidad de Los Andes, Chile; rmacera@uandes.cl

¶University of Chile, Dept of Industrial Engineering, Center for Applied Economics; cnoton@dii.uchile.cl

1 Introduction

Consider two hypothetical retail customers, Mr. A and Mr. B, in all identical except in the promotions they face. Mr. A is offered a deep discount for a given product whereas Mr. B is offered only a shallow one. Not surprisingly, Mr. A is more likely to buy the product than Mr. B. What happens, however, when these customers face subsequent promotions?

This paper investigates the dynamic effects of price promotions on future promotion response, i.e., the differences in behaviors between Mr. A and Mr. B after being exposed to different levels of promotional activity.[1] To this end, we implement a large-scale field experiment designed to identify dynamic promotional effects. Further, we put forward a search-related mechanism to rationalize our findings and derive the model's implications for the promotional activity of the firms.

How exposure to price promotions affects consumers' sensitivity to future promotions remains a controversial issue. On one hand, some managers' lay theories propose that having become accustomed to deep promotions, consumers become *less* sensitive to subsequent promotions, and therefore incentives need to be ratcheted up to induce the same incremental buying behavior. In this case, costumers are said to display *deal addiction*. On the other hand, it is also possible that deep promotions induce customers to become *more* sensitive to subsequent promotions, and therefore they require smaller promotional incentives after receiving large ones in order to maintain their purchasing behaviors. In this case, customers are said to display *heightened deal sensitivity*. The implications of these opposite views are extremely broad, as they affect firms' decisions about pricing (Anderson and Simester, 2004, 2010), about promotions and advertising (Erdem, Keane, and Sun, 2008), and may also affect firms' understanding of heterogeneity of consumer responses to several marketing mix decisions (Chan, Narasimhan, and Zhang, 2008).

We investigate the dynamic effects of price promotions through a large field experiment in the retail sector. In collaboration with one of the major retail chains in Chile, we exogenously vary the prices of 170 products across 17 categories and 10 retail stores to study how the depth of price discounts affects subsequent promotion sensitivities. To this end, we organize our experiment in two halves. In the first half (the intervention phase), products in 5 treated

[1]The marketing literature distinguishes between several price promotion instruments including temporary price reductions (TPRs) coupons, promotion packs, rebates, among others. We focus on TPRs –the most frequently used type of price promotion (Gedenk, Neslin, and Ailawadi (2010))– and use the term "price promotion" to refer to this specific price promotion instrument.

stores received a "deep" 30 percent discount while the same products received a "shallow" 10 percent discount in 5 similar control stores. In the second half (the measurement phase), the same products were attributed identical 10 percent discounts across all stores. We focus on systematic consumer behavior differences during the second half of the experiment, as a function of the intervention in the first half. This exogenous variation in promotion exposure, and the high-quality individual data (from the retailer's loyalty card program) allow us to evaluate the treatment effect, estimate a structural model to rationalize it and derive its supply-side implications.

Using standard regression analysis and a matching procedure, our experiment shows that deeper promotions make consumers more sensitive to future promotions. Customers exposed to 30 percent discounts during the first half of the experiment were 12.6 percent more likely to buy promoted items during the second half of the experiment than their control counterparts, despite promotion depths were held constant across all stores during the second half. Moreover, exposure to deep promotions increased the proportion of promoted goods relative to unpromoted ones in 15.8 percent. Estimates are robust to state dependence and are made conservative by the existence of stockpiling behavior.

In order to rationalize our heightened deal sensitivity result and derive supply-side implications, we use a model in which consumers search sequentially for promoted goods (Weitzman (1979)). The expected gains of searching are based on beliefs about discounts, which given the retailer's historical promotions, imply that deeper discounts today trigger larger expected gains of searching promoted goods tomorrow. As a consequence, the model predicts that consumers in treated stores should buy a larger share of promoted goods relative to their control counterparts. Regarding the supply-side implications, we find that small firms who are the least searched have a higher incentive to provide deep promotions in order to invite consumers to search. However, small firms are the most penalized once other competitors also offer deep promotions. In addition, we document that all competitors may become worse off when consumers take current promotions into account when predicting future deals.

This paper contributes to three streams of the Marketing literature. The first contribution of our paper is, to the best of our knowledge, to provide experimental evidence of the causal link between promotional activity and subsequent promotion sensitivity in a physical retail setting. The phenomenon of heightened deal sensitivity has been previously proposed in the retail context by Mela, Gupta, and Lehmann (1997) and Jedidi, Mela, and Gupta (1999).[2] Mela, Gupta, and Lehmann (1997) use a discrete choice model with time-varying parameters and document that, in the long-run, price promotions are associated with heightened price sensitivity of both loyal and non-loyal customers. Jedidi, Mela, and Gupta (1999) take advantage of the same long series analyzed by Mela, Gupta, and Lehmann (1997) and show that promotions are associated with negative brand equity.

Like ours, the work by Anderson and Simester (2004) also randomizes prices but in the context of mail order catalogs selling durable goods. They find that new customers who are offered deeper discounts exhibit larger orders in subsequent purchases relative to the control group. In contrast, established customers react in the opposite way by reducing their subsequent purchases. Our results complement those in Anderson and Simester (2004) as both, our context and experimental design, feature relevant differences. First, we analyze frequently purchased packaged goods in a supermarket environment, rather than durable goods in the mail order catalog context. This is important as some of the mechanisms explaining consumers' responses to promotion activity (e.g., learning) are likely to depend on the nature of the goods whose prices are being offered.[3] Second, our focus is on how promotion sensitivity is affected by the depth of past promotions, rather than on how overall purchase levels or subsequent regular price elasticities are affected by the depth of past promotions. In contrast, Anderson and Simester (2004)'s main focus is on subsequent reactions to price rather than to promotions.[4]

The second research stream we contribute to is the supply-side effects of promotions, by investigating the implications of heightened promotion sensitivity on promotional activities and equilibrium profits of competing manufacturers.[5] Our model suggests that small firms have incentives to provide deep promotions. However, rather than being able to take ad-

[2]See Neslin and Van Heerde (2008) for an extensive review on promotion dynamics.

[3]Because we manipulated promotional prices of well-known products, we believe the learning explanation is likely to play a less important role in our setting. See Tuchman, Nair, and Gardete (2017) for a similar argument.

[4]Even though Anderson and Simester (2004) also investigate deal sensitivity, they acknowledge that their design is not particularly well suited for this purpose, as they did not explicitly manipulate the timing of subsequent promotions. In effect, Anderson and Simester (2004) suggest our focus as a possible avenue for future research: "We cannot say how customers would have responded to [..] a subsequent discount. Investigating these issues would require different studies in which the experimental manipulations were [..] repeated in a subsequent catalog."

[5]See Raju (1992), Lal and Matutes (1994), Lal and Matutes (1994), Freimer and Horsky (2008), Villas-Boas and Villas-Boas (2008) for work considering the supply-side dynamics of price promotions, among others.

vantage of heightened promotion sensitivity in future periods, competing firms may end up worse off due to deeper discounts in a more competitive equilibrium. This is an example of the so-called Bertrand supertraps, as proposed by Cabral and Villas-Boas (2005), in which an apparent advantage for a monopolist (such as scope economies) can lead to all-across lower profits to competitors (through fiercer competition). Our model, therefore, breaks from the standard emphasis of the search literature, which has focused on consumer behavior by analyzing how search can affect the supply side.

Finally, our paper also makes a contribution to the estimation of consumer search models. We build on the seminal paper by Kim, Albuquerque, and Bronnenberg (2010) who estimate the theoretical search model of Weitzman (1979), which has been a go-to framework to analyze consumer search with limited information as in Honka and Chintagunta (2017), Chen and Yao (2017), Kim, Albuquerque, and Bronnenberg (2017).[6] To calculate the alternative-specific "reservation values" needed to simulate each consumer's search paths across products, Kim, Albuquerque, and Bronnenberg (2010) propose an interpolation table mapping consumers' product valuations to a set of pre-calculated reservation values. Instead, we propose two new alternative methods. First, under Logistic preference shocks, we show that the reservation values can be directly obtainable through a closed-form expression; and second, we prove a contraction mapping result for a broad class of continuous distributions that allows a fast iterative procedure to estimate reservation values.

The remainder of the paper is organized as follows. Sections 2 describes the experimental design, the data and verify the experimental manipulation. Section 3 shows the results of the field experiment. Section 4 motivates and describes the sequential search mechanism we propose, while Section 5 describes the model estimation and study the supply-side implications for promotional activities. Section 6 concludes.

2 The Field Experiment

This section describes the design and implementation of our experiment developed to study how consumers, who faced different discounts in the past, respond to subsequent promotion stimuli of the same magnitude. Subsection 2.1 describes the retail environment, the category

[6] See also Seiler (2013), Honka (2014), Bronnenberg, Kim, and Mela (2016) and Ke, Shen, and Villas-Boas (2016) for recent theoretical and empirical advances in modeling and understanding consumer search.

and product selection, the timing and structure of our intervention, and the randomization procedure. Subsection 2.2 describes the data made available by the retailer, while Subsection 2.3 assesses compliance of the intervention to the design by comparing actual and scheduled prices. We conclude that the implementation was successful and identify a few less successful categories.

2.1 The Experimental Design

To conduct the field experiment, we partnered with one of the two largest supermarkets chains in Chile commanding a market share in the order of 30 percent nationwide. The experimental intervention involved manipulating prices across 12 retailer stores, approximately one third of the number of stores operated by the retailer at the time.[7] In particular, the intervention involved assigning timing and depth of promotions to 170 top-sale products belonging to 17 categories: Beer, Bread, Breakfast Cereal, Candy, Cheese, Cold Cuts, Cookies, Cooking Oil, Fruit Juices, Meats, Milk, Pasta, Snacks, Soft Drinks, Tea, Water and Yogurt.[8] Within each category, and in order to maximize the "visibility" of the intervention, we randomly selected 10 products (or sku's) from the subset of 15 products exhibiting the largest market shares in each category.[9]

In order to analyze the intertemporal response to varying promotion depth, our experiment considers a total of 10 experimental weeks divided into two halves (or phases) of equal length. We use a coin flip to randomly assign stores to one of two alternative conditions: (i) In "treated" stores, participating products featured deep (30 percent) discounts during the intervention phase (five weeks), and shallow (10 percent) discounts during the measurement phase (five weeks); (ii) In "control" stores, participating products featured shallow (10 percent) discounts during both experimental phases.[10]

[7] The retailer's stores are organized into two retail sub-chains, which make use of different branding and perform separate marketing activities. While we performed the same experimental manipulation in both retail chains, we only report the results for the first chain, which is the larger one. Although the intervention failed to produce statistically significant results for stores of the smaller chain, importantly, all results are directionally in line across chains. We believe the smaller sizes of stores in chain 2 are partially responsible the absence of statistically significant effects. The results of chain 2 are available from the authors upon request.

[8] We detail the criteria used for category selection in Appendix A.

[9] Sku, or *stock keeping unit,* is a unique identifier of the product at the retailer.

[10] Regular prices at treated and control stores are very similar across store pairs, since the retailer sets regular prices based on "pricing zones". Pricing zones are defined by the retailer based on geographic location, demographics and competition intensity. Each pair of treated/control stores always belongs to the

The discounts of 30 and 10 percent are near the upper and lower bounds of the range of discounts typically used by the retailer. They were set as close as possible to these bounds in order to maximize the visibility of the intervention. In addition, we implemented a strictly positive low discount in order to isolate pure promotion effects from regular price effects.

Following the retailer's standard practice, each product remained in the promotion condition for one week. Products were placed on promotion on a Tuesday and remained in that condition until Monday of the following week. We rotated the products to be promoted within each category weekly, in order to keep promotion frequencies consistent with the retailer's practice. This process is illustrated in Figure 2. Each color/pattern in Figure 2 represents a different pair of sku's within a given category. In each of the first five weeks of the experiment the price of two different sku's were marked down (30 percent in treated stores, 10 percent in control stores). By the end of week five, the prices of all participating sku's had been intervened with. This process was repeated during the measurement phase, this time with equal mark downs of 10 percent in both types of stores. As we detail in Section 2.3, logistical reasons led us to opt for an approximate schedule to the one depicted in Figure 2 during the second half of the experiment, in order to take the managers' store-specific logistical constraints into account.[11]

We rely on a pairwise randomization over similar stores along three dimensions:[12] (i) similar consumer demographics (age, gender, socioeconomic group); (ii) similar competition intensity; (iii) similar geographic location. To classify stores according to their costumer demographic similarity, we used historical scanner data from the retailer's loyalty card program (see details in Subsection 2.2). Besides contributing to the pairwise randomization, using individual level data has several advantages. First, they allow us to reduce experimental noise by including pre-experimental behavior-based measures in the analysis. These types of measures are well known to outperform other context-independent measures such as

same pricing zone.

[11]We were able to get buy-in for our manipulation during the first half of the experiment. However, the retailer saw the promotion schedule as too heavy-handed for specific store-category combinations, and so an approximate alternative promotion schedule was agreed upon for the second half. The reasons for the changes were mainly due to previous agreements with manufacturers regarding some target of sales and the planned promotion activity for our promoted products and their respective competitors. The changes were at a national level and were not related to demand conditions. Moreover, they are constant across control/treated store pairs. As a result, they are unlikely to affect the validity of our estimates.

[12]In a pairwise randomization the full sample is divided into pairs, on the basis of covariate values, and within each pair one member is selected randomly to be assigned to the treatment group and the other member of the pair is assigned to the control group. See Imbens (2011).

demographic characteristics. Second, they allow us to verify the experimental manipulation ex post. Regarding the competition intensity, we follow the retailer's classification of each store into one of three levels of competition depending on the number and type of rival stores located within a certain radius of its stores. Regarding the geographic similarity, we aim at accounting for local unobservables such as weather shocks, demand conditions, etc.

2.2 Data

We implemented the experiment between the months of August and October of 2013. Even though 12 stores were originally selected to participate, only 10 stores actually participated in the experiment because the management decided to withdraw two of them after one week into the experiment (one assigned to the treatment group and one to the control group).[13]

The scanner data provided by the retailer cover all transactions carried out in the participating stores within the experimental period and 46 weeks before the intervention. The data include quantities purchased and the actual prices paid by a given customer for each sku in all product categories (including both intervened and not intervened categories). We are able to track the purchases of a given consumer using identifiers from the retailer's loyalty program database. This loyalty program covers a substantial fraction of the retailer's total revenues (approximately 80 percent). As mentioned before, we were also provided with cardholders' demographic information including their gender, age and socio-economic group classification. Our customer-level data spanning the experimental period include purchase information for 338,687 customers who visit our 10 stores, on average, every 2.4 weeks. Table 2 shows the descriptive statistics of the data per store-week pair. On average, each store had about 15,000 customer visits each week, selling an average of approximately 300,000 sku's weekly.[14]

Since treatment occurs in the first half of the experiment, we define the subpopulation of customers who took part in the experiment as those who visited a participating store and purchased any sku— in an intervened product category or not —anytime during the first half of the experiment. In order to avoid our results from being contaminated by customers

[13]The stores that were excluded from the experiment had originally been assigned to different store pairs. However, the fact that they shared the same classification based on customer demographics and competition level allowed us to reassign their couples to a new store pair without compromising our experimental design.

[14]For "bulk" purchases such as meat, for example, items are counted according to the number of packages bought rather than to the total quantity bought, typically priced by weight.

visiting different stores during the first half of the experiment, we excluded from our final sample those who switched stores in this intervention phase. Our focus on customers who visited and purchased any sku during the intervention phase of the experiment and did not switch stores over this period reduces the total number of customers in the sample to 286,162.

We use the 46-week pre-experimental data to construct measures of historical shopping behavior such as the total weekly expenditure, expenditure on intervened categories, and the average interpurchase time. This dataset includes 204,992 customers who are also in our experimental dataset (i.e., they visited a store and made a purchase during the intervention phase of the experiment). We exclude new customers (i.e., those for whom we lack historical data) from our working sample. Finally, we drop 83,755 customers whose pattern of store visitation departs significantly from that of the average customer (i.e., either visits a store too frequently or too infrequently). Our final sample includes 121,237 customers who visited and made a purchase at a participating store during the intervention phase of the experiment, did not switch stores over this experimental phase, and did not exhibit an abnormal visitation pattern during the experimental or pre-experimental periods.

Figure 3 compares baseline shopping behavior across experimental conditions. More specifically, it compares the pre-experimental distributions of two measures of customer purchasing behavior: total expenditure per visit and expenditure on promoted categories per visit. We observe a high degree of overlap between the distributions of treated and control customers for both measures of pre-experimental shopping behavior. Furthermore, average pre-experimental expenditure per visit in both promoted and nonpromoted categories appear to be very similar across experimental conditions as can be gauged from the proximity between the blue and red vertical lines depicting average expenditure for treated and control customers, respectively.

Table 2 presents further details on the comparison between treated and control customers. Individuals in both experimental conditions are highly similar along demographic dimensions and pre-experimental shopping behavior. A comparison of the demographic profile of treated and control groups shows that both groups are similar in terms of gender (a 63.9 percent of customers is female in the treated group versus 63 percent in the control group) and average age (46.8 years old in the treated group versus 45.1 years old in the control group). Complementing the information provided by Figure 3, Table 2 shows that differences in pre-experimental expenditure in promoted and nonpromoted categories are slim: Customers in the treated condition spend an average of \$43.1 dollars per visit (\$3.2 in promoted categories) versus an average of \$42.1 dollars spent by customers in the control condition per visit (\$3.3 in promoted categories).

2.3 Compliance With the Experimental Design

We partnered with the retailer to ensure the experiment was implemented according to plan. We first held a meeting with all the store managers to explain the importance of following the experimental protocol closely and maintaining an identical shopping environment in treated and control stores. Given that managers are often limited in the extent of the promotions they can offer, the experiment was generally well received even by managers of control stores who were allowed to discount products by only 10 percent. Second, an executive from the retailer headquarters was named as a coordinator and supervisor of the experiment. The coordinator was in charge of ensuring that price lists were sent to the participaring stores in a timely fashion. Third, store managers were asked to write back to headquarters and submit photos of updated price promotions on a weekly basis, as depicted in Figure 4.

To assess compliance, we compare the average price levels that were effectively implemented against the price dictated by the experimental design across treated and control stores. We perform the comparison for the implementation and measurement phases broken down by weeks and promotional status of the products.

Table 4 shows the comparison of the average prices during the intervention phase (first half) for the 17 categories involved in the experiment. The price comparison for the unpromoted and promoted products are in the left and right part of the table, respectively. As planned, the average price difference of unpromoted products between treatment and control stores is -0.2 percent, close to the ideal of the experimental design. Similarly, the average discount of promoted products was 9.7 and 26.6 percent in treated and control stores, respectively, very close to the ideal intervention of 30 and 10 percent discounts. Thus, the overall implementation was successful during the first half. Nevertheless, 4 shows that the Candy category was promoted at 50 percent discount levels in both treated and control stores. While this variation may still be exogenous, it does not inform us of the main research question, which is to analyze dynamic effects of shallow vs. deep promotions.

Table 5 shows the comparison of the average prices during the measurement phase (second half) for the 17 categories involved in the experiment. As in the implementation phase, there is little difference between prices across control and treated stores for the unpromoted products (-0.6 percent). Similarly, the average discount of promoted products were of 11.2 percent and 11.0 percent in treated and control stores, respectively. This indicates that the implementation during the measurement phase was also successful. However, during the measurement phase, the level of the discounts are extremely high for the Candy and Cheese categories, and actually negative for the Cooking Oil category. As discussed above, even if exogenous, such variation does not inform us of how consumers react to shallow/regular-sized promotions after having been exposed to different promotional conditions.

Finally, we also compare prices across experimental halves. Table 6 shows the comparison of the average prices between the intervention and measurement phases for the 17 categories. The price comparison for the unpromoted and promoted products are in the left and right part of the table, respectively. The first two columns of Table 6 reveal that regular prices increased slightly across the experimental halves, with average increases of 2.3 percent and 1.8 percent in control and treated stores, respectively. Within-category differences fall within single digits, and as a result, we see little reason to worry about any particular category based on these figures. The third column, however, reveals that promoted prices in control stores changed quite dramatically for the Cheese, Cooking Oil and Meats categories. The *expected* change is equal to zero because products in control stores are supposed to always be promoted at the 10 percent level throughout the experiment. Moreover, in the second half of the experiment, products in treated stores should also be promoted at 10 percent. The fourth column of Table 6 calculates the relative difference between promoted prices between treated stores in the second half, and control stores in the first half. As with the third column, the same three categories register high variation.

Despite the overall compliance with the experimental design, from the analysis above we identify four category exceptions that does not meet the scheduled discounts. The manipulation checks above reveal that the Candy, Cheese, Cooking Oil and Meats category prices are inconsistent with the experimental protocol. Even if the discrepancies are exogenous, their occurrence confuses the identification of the experimental question. In order to ensure that the experiment is answering our research question, we eliminated these four problematic categories from the analysis.[15]

[15] In terms of the implications of eliminating these categories, all of the results of interest survive their inclusion. However, a few additional confounding results arise, related to unexpected significance of a few variables with no apparent explainable pattern. These inconsistencies constitute another reason to focus the

3 Field Experiment Results

This section presents the field experiment evidence of the effects of promotion depth on the responses to subsequent promotions. Subsection 3.1 presents the OLS results using the full sample of customers, which suggest that the exposure to deeper promotions cause heightened deal sensitivity. To improve efficiency, Subsection 3.2 complements our experiment with a matching procedure that yields matched subsample with identical pre-experimental behavior at the individual level. Subsection 3.3 presents the OLS regression results using the matched sample, which confirm our heightened deal sensitivity finding.[16]

3.1 Full Sample Analysis

To ensure that only dynamic effects are captured, we focus on mean differences in consumer behavior between the control and treated stores during the measurement phase (second half) of the experiment. We run the following OLS regression,

$$y_i = \alpha + \beta_1 T_i + \beta_2 X_i + \varepsilon_i \tag{1}$$

where y_i is one of the following dependent variables for consumer i : i) whether the costumer is more likely to buy a promoted item; ii) the proportions (quantity and expenditure) of promoted vs. unpromoted items bought; iii) quantities and expenditures on promoted and unpromoted items. The main explanatory variable is T_i, an indicator function that equals one for customers of treated stores thus capturing the event that a customer was exposed to promotions with 30 percent (instead of 10 percent) discounts during the first half of the experiment. Given the randomization, the exogeneity of treatment T_i is ensured. We include controls X_i for precision and robustness purposes, namely store-pair fixed effects and individual-level controls such as customer gender, age and income levels, all in the form of indicator variables.[17]

The standar error estimates of equation (1) should account for two issues: heterocedasticity and the small number of clusters. First, since consumers of a given store may share

analysis on the unproblematic categories.

[16]Appendix B presents model-free effects of the intervention as complementary analysis to the formal econometric regressions. Results are consistent with heightened deal sensitivity.

[17]The age indicator variable is defined in spans of 20 years, and income levels represent socioeconomic groups as coded by the retail chain.

observable and unobservable characteristics, each consumer's effective contribution to reducing standard error estimates is lower than in the i.i.d. case. We address this potential heteroscedasticity issue by clustering our standard errors at the store level in order to avoid downward biases in to our standard error estimates (see Bertrand, Duflo, and Mullainathan (2004) for an in-depth discussion).[18] Second, the validity of standard error clustering relies on the asymptotic behavior of the estimator at the cluster rather than at the individual level. To account for the fact that we have a small of clusters/stores, we introduce a finite correction to the standard error of the treatment effect. We implement the "cluster residual bootstrap-t" procedure, as proposed by Cameron, Gelbach, and Miller (2008), to correct for downward bias related to small samples.[19]

The OLS results using the full sample suggest that consumers exhibit heightened promotion sensitivity after being exposed to deeper promotions. Table 7 summarizes the estimates of Equation (1). Columns (1) to (7) show the different dependent variables, while rows (a)-(c) show the p-values of the estimated treatment effects using OLS standard errors, clustered standard errors and the "cluster residual bootstrap-t" as discussed above. Column (1) shows that treatment induces an increase in the probability of buying a promoted sku in 4 percentage points. Using the statistic $E\left[\widehat{y}_i \middle| X_i, T_i = 1\right] \div E\left[\widehat{y}_i \middle| X_i, T_i = 0\right]$ (where the expectation is over the sample of consumers), we find an implied relative increase of 10.5 percent. This dynamic effect is reasonable, given that retail settings are extremely busy with Marketing stimuli. Columns (2) and (3) show that consumers also have a higher share of promoted items in their baskets during the measurement phase of the experiment. Estimates in columns (1) to (3) are statistically significant across standard error specifications with p-values below 0.06. Finally, columns (4) to (7) suggest that treated consumers also exhibit higher purchase rates of all items, independently of promotional condition. However, these results are only significant when OLS standard errors are used.

[18]This approach is sometimes seen as too conservative, potentially increasing the probability of type 2 errors. Given the relative similarities of store pairs in our experiment it is possible that, in our case, the correct specification would produce standard error estimates in between the ones provided by OLS and by clustering at the store level.

[19]See Cameron and Miller (2015) and Imbens and Kolesar (2015) for relevant discussions and practical guidance.

3.2 Improving Efficiency Through Matching

To improve the efficiency of our treatment estimates, we complement our field experiment with a matching procedure. Much like the use of control regressors, a matching procedure reduces the variance of the unobserved error term by taking advantage of the correlation between observable and unobservable characteristics. This approach shifts the emphasis from the cluster to the individual level by matching, within pairwise randomized stores, those individuals who have balanced covariates before the experiment (Rubin (1973, 1979); Imbens and Rubin (2015)).

In our case, the matching technique allow us to construct a large subsample of *statistical twins* (one twin buying at a control store and the other at a treated one) to ensure identical pre-treatment purchasing behavior between treatment and control *at the individual rather than at the store level.*.[20] Notice, thus, that in our setting the exogeneity of the treatment is guaranteed by the experimental design and matching is only needed to identify similar customers based on historical data.[21] As we discuss in the result section below, main findings are not sensitive to the use of the matched sample, showing a treatment effect that is consistent across subsamples.

Matching Framework. To construct our sample of statistical twins, we introduce to the marketing literature a recent matching technique developed by Zubizarreta (2012). This matching technique takes advantage of new developments in optimization to match individuals in multiple dimensions, an until recently unfeasible task due to the dimensionality of the problem. Matching individuals on several dimensions encompasses other matching techniques, such as propensity score, by creating a superior and easily interpretable matching sample.

Formally, let $\mathcal{T} = \{t_1, \ldots, t_T\}$ be the set of treated units, and $\mathcal{C} = \{c_1, \ldots, c_C\}$, the set of potential controls. Without loss of generality, suppose $T \leq C$. Each treated unit $t \in \mathcal{T}$ has a P dimensional vector of observed covariates $\mathbf{x}_t = \{x_{t,1}, ..., x_{t,P}\}$, and each control $c \in \mathcal{C}$ has a similar vector $\mathbf{x}_c = \{x_{c,1}, ..., x_{c,P}\}$. Let the assignment indicator $a_{t,c}$ be equal to 1 if

[20] This matching procedure is standard when randomization is at the cluster level, but statistical analysis is at the individual level (Imbens, 2011).

[21] Unlike our paper, the matching technique is typically used with non-experimental data in order to balance the relevant pre-treatment covariates between treatment and control groups (Imbens and Rubin, 2015).

treated unit t is assigned to control c, and 0 otherwise; and denote the entire assignment matrix by $\mathbf{a}$.[22] The optimal assignment problem is given by:

$$\min_{\mathbf{a}} \sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} \delta_{t,c} a_{t,c} \tag{2}$$

$$\text{subject to} \sum_{c \in \mathcal{C}} a_{t,c} = 1 \quad , \quad t \in \mathcal{T} \tag{3}$$

$$\sum_{t \in \mathcal{T}} a_{t,c} \leq 1 \quad , \quad c \in \mathcal{C} \tag{4}$$

$$a_{t,c} \in \{0,1\} \quad , \quad t \in \mathcal{T}, c \in \mathcal{C} \tag{5}$$

$$\left| \sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} \frac{x_{c,j} a_{t,c}}{T} - \overline{x}_{\mathcal{T},j} \right| \leq \varepsilon_j \quad , \quad j \in \{1,..,P\} \tag{6}$$

where $\delta_{t,c} \in [0, \infty)$ is a distance function between treated and control units (e.g. Euclidean distance), $\sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} \frac{x_{c,j} a_{t,c}}{T}$ denotes the average covariate j of assigned controls and $\overline{x}_{\mathcal{T},j}$ denotes the average covariate j across all treated individuals.

The goal of the matching program is to minimize the total sum of distances between treated units and matched controls as stated in expression (2). The first three constraints describe the integer nature of the assignment problem: Each treatment unit is paired with one control unit (Equation (3)) and not all control units should be used (Equation (4)).[23] The set of constraints given by expression Equation (6) introduces an upper bound on the difference allowed between treatment and control individuals for each covariate, according to $\varepsilon_j > 0$, the pre-determined tolerance level for covariate $j \in \{1,..,P\}$. This last set of constraints is a distinctive feature of themixed-integer programming (MIP) matching approach proposed by Zubizarreta (2012).

Matching Supermarket Customers. We use the individual-level data available through the retailer's loyalty card to match pairs of control and treated consumers based on the pre-experimental records according to the procedure described above. To construct statistical twins, we consider demographic and behavior-based covariates: Age, gender, weekly average of total expenditure, weekly average of total expenditure in the experimental categories, and the frequency of trips to the store. The historical dataset used for this task covers a 46 week

[22]The assignment problem is a mixed-integer programming (MIP) problem where some of the decision variables are constrained to be integer values at the optimal solution.

[23]Note that, for the assignment problem, the labeling of treatment and control units is irrelevant for its optimal solution.We relabel some stores to increase the sample size of the matched sample.

period, one year prior to the experiment. We focus the analysis on consumers who visited a store at least once during each experimental phase.

Table 8 presents the sample sizes of the universe of customers before matching and those who were matched by the MIP matching procedure. Columns (1) and (2) present the number of customers who faced the experimental promotional activity in each store pair and for whom we have historical data. Overall, as shown in Columns (3), (4) and (5), the MIP matching generated $31,586$ one-to-one customer pairs (one control and one treated customer), distributed across 10 stores of our retail chain. Columns (6), (7) and (8) presents the final sample of $14,465$ matched pairs of individuals.

Table 9 presents the resulting covariates of the final matched sample.[24] The last column reports the p-value for the null hypothesis of identical means. We obtain near identical means in total expenditure, expenditure in promoted categories, and age between treatment and control matched individuals. Given the large sample size of individuals, the tests reject nearly identical means in gender and number of trips although the table shows that the actual values are quite similar.

Importantly, the matching procedure was not revisited after the final tolerance levels were established, and all analysis was performed after the completion of the matching procedure. This sequence of events ensures that the matching procedure is not contaminated by the results it generates, eliminating the potential for feedback effects and researcher bias.

3.3 Matched Sample Analysis

We re-estimate Equation (1) in Subsection 3.3 using the customers who were matched across treatment and control conditions using the MIP procedure. As a reminder, we focus on customer behavior during the second half of the experiment in which promotion depths are held constant across experimental conditions.[25]

Table 10 presents the results using the matched subsample to confirm that consumers

[24] We explored different values for the tolerance parameters ε_j for each covariate $j \in \{1, .., P\}$, to account for the trade-off between the proximity measures of the paired customers and the resulting sample size. On one hand, large values of ε_j lead to poorly matched pairs, while on the other, smaller values of ε_j reduce the sample size. In fact, some combinations of small values of ε_j imply no feasible solutions, i.e., no assignment meets the desired levels of balance on covariates.

[25] Clustering standard errors is also important when using the matched sample to take into account the fact that customers of the same store may be exposed to correlated unobservable shocks. Standard errors are clustered at the store level and p-values for the treatment effect are derived from the cluster residual bootstrap-t procedure with 50,000 draws.

display heightened deal sensitivity. The first column presents the effect of deeper promotions on the purchase incidence of promoted products as the dependent variable. Treatment induces an increase in probability of buying a promoted sku in 5.6 percentage points, and is statistically significant at the 95 percent confidence level. The statistic $E\left[\widehat{y}_i \middle| X_i, T_i = 1\right] \div E\left[\widehat{y}_i \middle| X_i, T_i = 0\right]$ reveals consumers exposed to deep promotions are 12.6 percent more likely to buy promoted items than their control counterparts, despite facing similar promotional levels. As before, the second and third columns suggest a shift in consumers' basket compositions towards promoted products. The proportion of promoted goods in consumers' baskets increased by 15.8 percent in relative terms, and the share of wallet increased in 15.9 percent.

The point estimates in Table 10, columns (4)-(7), further suggest that the shift in basket compositions appears to source from an overall increase in the total number of purchases, with a disproportionate increase of promoted products vs. unpromoted ones. However, we prefer not to emphasize these results, given the lack of statistical significance of the treatment effects in these columns.

The heightened deal sensitivity result holds across categories. To show this, Figure 5 summarizes the category-specific estimates of the probability of buying a promoted item. The point estimates range from 0.032 for the Cold Cuts category to 0.102 for Breakfast Cereal. More importantly, all category estimates are positive, and moreover almost all treatment effects are significant at the 5 percent level. As such, we find no evidence in support of the alternative hypothesis that deeper promotions lead to consumer deal addiction.

As a robustness check, we compare the point estimates using the full and the matched samples. As discussed in Section 3, the matching procedure was implemented to reduce estimation noise. However, the matched sample may have also focused the analysis on a subset of consumer with specific treatment effects, different from the ones exhibited in the population. Comparing the point estimates in Table 7 to those in Table 10 , we verify that, except for the coefficient of expenditure on unpromoted products, all coefficients from the regression using the full sample fall within the confidence intervals of the coefficients using the matched sample. This is suggestive evidence that the matching procedure did not alter the distribution of treatment effects in a statistically significant way. However, it is possible that the matched sample exhibits a lower increase of purchase rates of unpromoted products, as a result of differential exposure to promotions. We cannot reject the hypotheses that the remaining coefficients are identical.[26]

Appendix C considers a number of further robustness checks. First, we consider different econometric specifications for each of the dependent variables, in order to take their bounded nature into account. This analysis suggests that our use of linear specifications underestimates the statistical significance of treatment effects. Second, whenever we consider a placebo test that performs the same previous analysis but with pairs of customers who did not visit the retailer stores during the first half of the experiment—and thus were unlikely exposed to the intervention—we find no statistical evidence of heightened promotion sensitivity, as expected.[27]

4 Mechanism: Searching for Deals

Our experimental findings suggest that the heightened promotion sensitivity caused by deeper promotions is economically significant. We use a structural model of consumer behavior, specifying a well-defined search mechanism that can rationalize our findings, to derive supply-side implications. Subsection 4.1 presents reduced form evidence of the search behavior in our data and discusses alternative mechanisms. Subsection 4.2 presents the structural model of search-behavior.

4.1 Reduced-Form Evidence of Consumer Search and Alternative Mechanisms

The role of consumer search in retail markets is widely accepted, and has been the focus of a long list of empirical applications, either directly or sometimes as a benchmark for comparison with online settings (e.g., Warner and Barsky (1995), Urbany, Dickson, and Kalapurakal

[26]We have simplified the comparison of coefficients by conditioning on the treatment effects from the original analysis with no matching. This method is used as an approximation, since the full analysis is complicated by the fact that the matched sample is necessarily correlated with the original one. Theoretically, without information about the joint distribution of the matched and unmatched coefficients, an implementation of non-nested tests à la Vuong (1989) is unavailable. While it is theoretically possible to use a bootstrap approach that performs the MIP matching procedure on each bootstrap sample of the original dataset, the additional requirement of nesting the finite bootstrap-t small sample correction implies an impractical amount of computation time.

[27]Results on the heterogeneity of treatment effects are available from the authors. We find treatment effects are positively correlated with the number of promoted items bought in the past, controlling for average basket sizes. This is consistent with the idea that consumers who value promotions more, or equivalently face lower costs when searching for deals, also show the largest increases in promotion sensitivities.

(1996), Brynjolfsson and Smith (2000), Clay, Krishnan, and Wolff (2001), Goolsbee (2001), Erdem, Imai, and Keane (2003) and Seiler (2013)). We consider the conditions for this explanation to rationalize our findings and we discuss other alternative mechanisms that may also be in play.

Reduced-Form Evidence of Search. We explore the link between consumer search and the heightened promotion sensitivity caused by the deeper promotions. Since promotional prices are not easily observable in the retail setting, consumers are likely to search for deals if they believe there to be attractive promotions. As a result, consumer search will explain the heightened promotion sensitivity if our treatment can generate larger expected gains of searching for promoted goods.

By investigating the historical promotional patterns at the category level, we find that in general, promotions depths are positively correlated across consumer visits to our retailer. The transition matrix of promotions presented in Table 13 shows that the probability of facing a deep promotion today is higher when deep promotions were available during the previous visit, the same applying to shallow promotions for most categories.[28] This indicates that rational consumers, who form beliefs that are consistent with the historical promotional patterns of the retailer, must believe that deep promotions are more likely if they were exposed to deep promotions in the past, a necessary condition for our search model to rationalize our heightened promotion sensitivity result.

Additional evidence follows from the fact that categories with stronger positive autocorrelation of promotion activity display larger treatment effects. We regress the category-specific treatment effects on the lift in the probability of deep promotions, $Pr\left(Deep_t \middle| Deep_{t-1}\right) - Pr\left(Deep_t \middle| Shallow_{t-1}\right)$, as shown in Figure XXX. If consumer beliefs are consistent with empirical promotion patterns, then we should observe a positive association between our treatment effects and difference in likelihood of facing deep promotions. The correlation is of 22 percent, and increases to 57 percent when non-statistically significant categories at the 5 percent level are removed from the analysis (see Figure XXX).[29]

[28] We discuss the estimation procedure of the promotional processes in Section 4.2.

[29] We exclude the Soft Drinks and Yogurt categories, in which deep promotions are altogether extremely rare and also the shallow promotion state is an absorbing one for Soft Drinks.

Alternative Mechanisms. We discuss and test for alternative mechanisms that can also be present in our setting such as forward buying and state dependence. We conclude that our search mechanim remains as an important—though non necessarily exclusive—mechanism for our findings.

First, forward buying/stockpiling behavior has been proposed as one of the forces affecting promotion sensitivity (see Blattberg and Neslin (1993), Mela, Jedidi, and Bowman (1998) and Anderson and Simester, 2004 for insightful discussions). Our experiment is likely to affect consumer stockpiling behavior, given the differential discounts offered during the intervention phase. For example, treated customers may take advantage of deep discounts to stockpile some products to a higher extent than their control counterparts. However, stockpiling behavior works counter finding heightened promotion sensitivity, since it would theoretically induce treated consumers to react to subsequent promotions less than their control counterparts because of the larger inventories held at home. Hence, stockpiling behavior is unlikely to underlie our results and, if present, underestimates our finding of heightened promotion sensitivity.

Second, several dynamic mechanisms predict consumers are more likely to repeated purchases of products they have bought in the past. We refer to this well-established empirical phenomenon as state dependence, which was operationalized by Guadagni and Little (1983) and further analyzed in detail by Dube, Hitsch, and Rossi (2010). This phenomenon can be generated by several psychological mechanisms, including consumers' thinking costs (Shugan, 1980) and consumer consideration sets (Hauser and Wernerfelt, 1990), among others. To understand how state dependence could explain our results, consider a pair of similar customers who only differ on the experimental condition they were exposed to in the intervention phase. Assume that the treated customer bought a promoted product because of the deep discount, whereas the control customer decided not to purchase it, given its lower promotional discount (10 percent). It is possible that, during the second half of the experiment, both customers visited the store on the week that the same product is on promotion again (albeit at a shallow level). In this case the treated customer may be more likely to buy the product in the second half not because of heightened promotion sensitivity, but rather because of state dependence, triggered by the deep discount in the first experimental half. Therefore, any mechanism driving state dependence is necessarily product specific as consumers must purchase the product to create a dependence with it.

Using the fact that state dependence is product-specific, we repeat our main analysis, but now eliminating purchases of goods that were also bought during the first half of the experiment. Table 11 presents the differential effects of deep price promotions on subsequent sensitivity to promotions of products that were not previously bought. The results reveal that exposure to deep discounts makes treated customers more likely to buy items on promotion later on, even if they did not buy them during the first half. The estimates of interest change little when compared to the ones in Table 10, and the significance of the treatment effects across the first three columns increases. The absence of sign and magnitude differences with the results of Table 11 confirms that state dependence is unlikely to be the main responsible for the estimated treatment effects.

4.2 Model

We now present a search model that rationalizes our experimental findings and use it for counterfactual analysis. The model captures the intuition that our manipulation led consumers in the treated group to expect higher discounts during the second half of the experiment, and thus to prioritize more the evaluation of promoted goods in hopes of finding future good deals relative to the control group.

Optimal Search Rules. We adopt the model of sequential search of Weitzman (1979) in which rational forward-looking consumers decide which products to evaluate, when to stop searching and which products to buy, if any. In the model, the consumer has to pay a cost $c > 0$ to observe the value of an unknown option. Since there are finite number of uncertain options, the consumer's problem is to find the optimal rules to (i) determine the order of the alternatives in her potential sequential search and (ii) decide whether to keep searching, to buy and stop, or not to buy and stop after the inspection of each option.

To illustrate, suppose a consumer decides whether to evaluate a product with unknown utility u_1, while having in hand an alternative good with known utility u_0, which remains available after the inspection of u_1. If the consumer searches, she expects to earn:

$$EUSearch_1 (u_0) = -c + Pr (u_1 \geq u_0) E [u_1 | u_1 \geq u_0] + Pr (u_1 < u_0) u_0 \tag{7}$$

which captures the fact that search may not necessarily be advantageous ex post, but depends on whether u_1 is higher or lower than the value of the current option, u_0. Notice that a consumer with a high value of u_0 is less likely to search than one with a low value, ceteris paribus.

Importantly, the optimal rules rely on a reservation value defined as the certain utility that makes the consumer indifferent between searching the next alternative or not. In our example, the reservation value of alternative 1, denoted by u_0^*, is implicitly defined by the solution to $EUSearch_1\left(u_0^*\right) = u_0^*$. Extending this idea to multiple options, the reservation value for product j, denoted z_j^*, satisfies:

$$z_j^* = -c + \int_{z_j^*}^{\infty} udF_j\left(u\right) + F_j\left(z_j^*\right)z_j^* \tag{8}$$

where $F_j\left(\cdot\right)$ is a product-specific cumulative distribution function that represents the beliefs about the uncertain utility of product j.

Weitzman (1979) shows that optimal sequential search with multiple alternatives is characterized by a simple two step rule: 1) calculate the reservation utility values z_j^* that make the consumer indifferent between searching each alternative j and not; 2) proceed to search alternatives by descending order of z_j^*, until the highest utility of the evaluated products is higher than all of the reservation values of the remaining uninspected alternatives.

Discount Depth Transition Probabilities. In order to compute the reservation values in Equation (8), we model consumer beliefs over discount depths through a Markov process that accounts for promotions that may not be independently distributed over time. Let ω denote the empirical probability that a focal category is promoted with a shallow discount, and $1-\omega$ the probability that the category is promoted with a deep discount. Since these probabilities may be time-dependent, we assume the promotion process follows a transition probability matrix featured in Table . Notation ω^S indicates the probability of a category

Table 1: Discount Depth Transition Probabilities of a Focal Category

		t:	
		Shallow Discount	Deep Discount
$t-1$:	Shallow Discount	$1-\omega^S$	ω^S
	Deep Discount	$1-\omega^D$	ω^D

being promoted with a deep discount in time period t, given that it was promoted with a shallow discount in period $t-1$. Similarly, ω^D denotes the probability that the category will be promoted with a deep discount given that it was promoted with a deep discount at time $t-1$. Notice that in this specification two consumers facing the same promotional activity may hold different beliefs over discount depths in period t because of exposure to different promotional activities during period $t-1$.

Consumer Utility. In order to compute the reservation values in Equation (8), we model consumer's utilities. Let consumer i derive utility from buying product j at purchase occasion t be equal to:

$$u_{ijt} = \alpha_j + \beta SDep_{ijt} + \gamma^D d_{jt}^{Deep} + \gamma^S d_{jt}^{Shallow} + \varepsilon_{ijt} \tag{9}$$

where α_j captures the time invariant *net utility* of buying the product and paying its regular price[30]; $SDep_{ijt}$ captures whether individual i purchased the product j in the previous period (state dependence); d_{jt}^{Deep} and $d_{jt}^{Shallow}$ indicate the discount depths of product j at time t; and ε_{ijt} is an individual taste shock that is an idiosyncratic fit of product j at time t.[31]

Through costly search, the consumer gathers information about the discount depth and the taste shock ε_{ijt}. We assume that consumers know the regular prices, which is consistent with our retail setting where regular prices rarely vary. Hence, when the consumer faces a product at a regular price, the taste shock is the only source of uncertainty. Consequently, Equation (9) for the utility of non-promoted products, denoted by u_{ijt}^r , reduces to:

$$u_{ijt}^r = v_{ijt} + \varepsilon_{ijt} \tag{10}$$

where $v_{ijt} \equiv \alpha_j + \beta SDep_{ijt}$ is the mean utility for individual i of product j in the absense of promotions.

Promoted products are featured with salient labels that invite consumers to inspect the discount offerings and assess whether the deals are attractive enough. The implication is that

[30]Because in our experiment regular prices were not manipulated and we observe little variation in historic regular prices, in the model we assume that regular prices do not change over time, $p_{jt}^r = p_j^r$. Therefore, $\alpha_j = \alpha_j^{'} - \gamma^R p_j^r$

[31]For example a consumer may want to deliberate on how much of a product she has at home before deciding whether to restock, or if a particular occasion will benefit from a particular meal, ingredient, etc (see for example Shugan (1980) and Hauser and Wernerfelt (1990)).

consumers can immediately identify the promotional status of products, but are required to inspect pricing labels to assess promotional values. Thus, when the consumer faces a promoted product, besides the taste shock there is additional uncertainty. Consequently, Equation (9) for the utility of promoted products, denoted by u^p_{ijt}, becomes:

$$u^p_{ijt} = v_{ijt} + \varepsilon_{ijt} + \begin{cases} \gamma^D & with\,probability\;\; \omega^H_{ijt} \\ \gamma^S & with\,probability\;\; 1 - \omega^H_{ijt} \end{cases} \tag{11}$$

where $H \in \{S, D\}$ is the promotional history, and ω^H_{ijt} is consumer i's transition probability that product j is promoted with deep discounts at time t, given history H.

Consumer Beliefs. In order to compute the reservation values in Equation (8), we model the consumer beliefs about the uncertain utility of product j, $F_j(\cdot)$. We assume $\varepsilon_{ijt} \sim \mathcal{L}(0, 1)$ where $\mathcal{L}$ is the logistic distribution, which largely simplifies solving the search problem as proved below. Consequently, the product utilities of regular priced and promoted products have the following distributions:

$$u^r_{ijt}\Big|_{v_{ijt}} \sim F^r = \mathcal{L}(v_{ijt}, 1) \tag{12}$$

$$u^p_{ijt}\Big|_{v_{ijt}} \sim F^p = \omega^H_{ijt}\mathcal{L}\left(v_{ijt} + \gamma^D, 1\right) + \left(1 - \omega^H_{ijt}\right)\mathcal{L}\left(v_{ijt} + \gamma^S, 1\right) \tag{13}$$

Equation (12) shows that the utility of non-promoted goods follows a logistic distribution consistent with most choice models. Equation (13) shows that the utility of promoted products follows a mixture of logistic distributions because in this case consumers not only face a preference shock, but they also face the additional uncertainty from the unknown promotion depth.

Calculation of Reservation Values. Our resulting distributions for the utilities of non-promoted and promoted products largely simplify solving the search problem, which requires finding the reservation values z^*_j defined by the non-linear equation (8). This task is time consuming when tackled through simulation as it must be solved for each iteration/consumer/product/period combination. Previous work overcome this issue by pre-computing lookup tables of z^*_j as a function of c and v_{ijt}. In addition, past work on empirical search estimation (e.g. Kim, Albuquerque, and Bronnenberg (2010), Honka and Chintagunta (2017)) has assumed normally distributed taste shocks, $\varepsilon_{ijt} \sim N(0,1)$. In constrast, the Proposition below shows that assuming logistically distributed taste shocks, $\varepsilon_{ijt} \sim \mathcal{L}(0,1)$, yields a simple closed-form solution to finding the reservation values for non-promoted goods.[32]

Proposition *(logistic uncertainty) The reservation value equation* (8) *admits a closed-form solution when $F_j(u)$ is logistic with location parameter v_{ijt} and scale equal to 1, namely $z_j^* = v_{ijt} - \ln(\exp(c) - 1)$.*

In our case, the result above is useful to recover the reservation values of regular-priced products, under the logistic assumption $\left. u_{ijt}^r \right|_{v_{ijt}} \sim \mathcal{L}(v_{ijt}, 1)$. However, it is not helpful to solve for the reservation values of promoted products, because in that case consumers also face uncertainty over promotion depths and the resulting distribution is a mixture of logistic distributions for which we have no closed-form solutions. We take advantage of the following contraction mapping result to compute the reservation values of promoted products:[33]

Theorem *(contraction mapping) Function $\Gamma(z) = -c + \int_z^\infty u dF_j(u) + F_j(z)z$ is a contraction mapping for any differentiable cumulative distribution function $F_j(z)$ with finite moments $E(u|\, u > z) \, \forall z \in \mathbb{R}$.*

This contraction mapping result can be used to efficiently and precisely recover the reservation values for a general class of distributions while controlling the precision levels of the solutions without the need for interpolation methods.[34] In particular, we can use this theorem to find the reservation values of promoted products that follow a mixture of logistic distributions as shown in Equation (13).

---

[32] All proofs are in Appendix D.1 and D.2.

[33] Let be a (X, d) metric space. Recall that a map $(T : X \to X)$ is called a contraction mapping on X if there exists $q \in [0, 1)$ such that : $d(T(x), T(y)) \leq qd(x, y)$ for all x, y in X. It guarantees the existence and uniqueness of fixed points of certain self-maps of metric spaces, and provides a constructive method to find those fixed points.

[34] This approach can also increase estimation speed slightly, but, depending on interpolation method, lookup tables can be extremely fast as well.

5 Model Estimation

In this section we investigate whether our field experiment findings can be rationalized by our sequential search model. In particular, we propose that the experiment changed beliefs about promotion depth, which in turn increased the search and purchase of promoted goods. In subsection 5.1 we describe the estimation procedure of our sequential search model. In subsection 5.2 we present the structural estimates for the milk category. In subsection 5.3 we study the supply-side implications of our findings.

5.1 Estimation Procedure

We are interested in estimating the probability of purchasing promoted goods within our sequential search model:

$$Pr\left(Choice_{ijt}|\theta, X_{ijt}, \omega_{it}\right) \tag{14}$$

where $Choice_{ijt}$ is the event that consumer i purchases product j at time t conditional on the vector of preference parameters $\theta = [\alpha_1, ..., \alpha_j, ..., \alpha_J, \beta, \gamma^D, \gamma^S]$ and the product characteristics $X_{ijt} = [SDep_{ijt}, d_{jt}^{Deep}, d_{jt}^{Shallow}]$ as introduced in Equation (9). Finally, ω_{it} is the consumer's belief about the promotional activity introduced in Equation (11).

Because ω_{it} is unobservable, we rely on the structural model to derive the search sequence of consumer i, S_{it}, given the optimal rule described in Subsection 4.2. Since search paths in retail settings are also unobservable, the probability of choosing a given alternative can be obtained by integrating over the search paths that led up to such choice:

$$Pr\left(Choice_{ijt}|\theta, X_{it}\right) = \sum_{S_{it}\in\mathcal{S}} Pr\left(Choice_{ijt}\left|S_{it}, \theta, X_{it}\right.\right) Pr\left(S_{it}\left|\theta, X_{it}\right.\right) \tag{15}$$

where $\mathcal{S}$ is the set of all potential search paths.

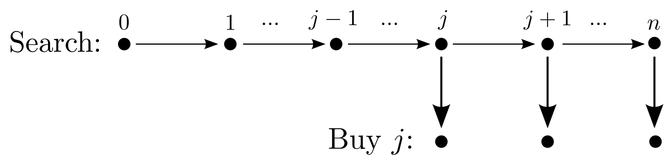
Computing Search Paths and Simulated Maximum Likelihood. For illustration purposes of the estimation process, consider a guess of preference parameters, θ_0, and assume only an outside option with known value v_0, and two products with unknown values v_1 and v_2 exist. Further, let the first product have a higher reservation value than the second, $z_1 > z_2$ for θ_0, consistent with Weitzman' rule. The consumer needs to compare the actual utility draws, v_j, in order to decide whether not to buy and keep searching, buy or not to buy at all. She opts for the outside option given each of the following search sequences:

$$\begin{array}{rc}
Path\,1:\ No\,search & \underbrace{v_0 > z_1}_{stop\,search} \\
Path\,2:\ Search\,1\,only & \underbrace{v_0 < z_1}_{search\,1} \wedge \underbrace{v_0 > v_1}_{prefer\,0} \wedge \underbrace{v_0 > z_2}_{stop\,search} \\
Path\,3:\ Search\,1\,\&\,2 & \underbrace{v_0 < z_1 \wedge v_0 > v_1}_{search\,1\,\&\,prefer\,0} \wedge \underbrace{v_0 < z_2 \wedge v_0 > v_2}_{search\,2\,\&\,prefer\,0}
\end{array}$$

Consequently, the consumer opts for the outside option with probability $Pr\left(Path\,1 \vee Path\,2 \vee Path\,3\right)$.

Analogously, the figure below shows the potential search paths consistent with the choice of alternative j.

Figure 1: Search Paths Consistent with Choice of j



In the diagram above, searching an additional option corresponds to a lateral movement, and a downward one depicts the purchase of alternative j. For a consumer to be willing to search option j with reservation value z_j, she must have inspected options with higher reservation values before and have found that it was worthwhile searching option j nonetheless. A consumer who did not search option $j-1$, will not search for option j either because options are ordered by their reservation values. After inspecting option j, the consumer may decide to stop search immediately, or instead may evaluate additional alternatives before deciding which product to purchase. In the appendix we provide a general expression for the choice probability of an arbitrary alternative j, which involves adding all of the possible paths.

Estimation of Equation (15) is performed through simulated maximum likelihood with two main challenges. First, a difficulty with *accept/reject* choice simulation is that small changes in parameters do not affect simulated outcomes, even for large sets of draws. Second, the likelihood function has a number of saddle points that make identifying the global maximum challenging. Both issues were solved by performing a global patterned search across parameter values. In addition, we employed the smoothed estimator proposed by McFadden (1989) to calculate the hessian of the log-likelihood and the resulting standard errors. We provide all the details in appendix D.3.

Identification of Preference Parameters. We now intuitively discuss the identification of vector of preference parameters, $\theta = [\alpha_1, ..., \alpha_j, ..., \alpha_J, \beta, \gamma^D, \gamma^S]$. We focus on the product-specific constants, α_j in Equation (9), as the identification of the remaining parameters follows the same argument. It suffices to note that when the alternative-specific constant of a parameter increases, it shifts *both* the probability of that alternative being evaluated and of being bought. The reason is that an increase of an alternative-specific constant increases the ex-post utility of that alternative as well as its expected utility, captured by its reservation value. However, given a search cost parameter c, each guess of alternative-specific constants corresponds to a unique set of reservation values, implying the identification of θ.

Identification of Search Costs. In the absence of search data, search cost cannot be identified as different combinations of c and θ can rationalize the same choice data. To see this, consider a market with two products, with market shares of 90 percent and 10 percent. Consumers may have an overwhelming preference for the first product because 1) it provides a much higher utility than the second product or 2) it is only slightly better than product two, but search costs are extremely high such that most consumers prefer not to consider the second product. As consequence, we have to normalize the search cost parameter, accounting for the fact that the search cost normalization is not always innocuous. In fact, while setting a low search cost allows the model to rationalize the data, a high value precludes the model from predicting close market shares across products: A gap in market shares of alternatives would necessarily emerge under large search costs, independently of the remaining parameters. No such limitations were found in our application.

5.2 Testing the Model

We now show that our structural estimates can explain the heightened promotion sensitivity result. First, we compute the transition probabilities of discount depth, ω, using pre-experimental data. Second, we feed the structural model using beliefs ω to estimate the preference parameters, θ. Finally, we test whether the size of the estimated treatment effect can be rationalized within our sequential search model.

Estimation of Discount Depth Transition Probabilities. We recover the discount depth distribution from pre-experimental data and assume that consumers have observed this history to build their beliefs. We assume the promotion process follows a transition probability matrix featured in Table 12.

We estimate the Markov process at the category level using standard count methods. We label a discount as shallow if the average discount for the category ranges between 0 and 15 percent and as deep if the average discount for the category is larger than 15 percent. In computing transitional probabilities, we assume an interpurchase time of 2 weeks, which accords well with the average time between visits that we observe in the data. Hence subscript t should be interpreted as purchase occasion from here on.[35]

We present the results of the discount-depth transition estimation in Table 13. With the exception of the Cookies category, consumers are most likely to find shallow discounts after being exposed to shallow discounts during their previous store visit. Moreover, deep discounts appear less "sticky" than shallow ones. Importantly, with the exception of the Soft Drinks category, consumers are generally more likely to find deep discounts during their current visit after being exposed to deep rather than shallow discounts during their last visit (i.e. column 4 dominates column 2). This is consistent with the search explanation that consumers may be willing to evaluate more discounted alternatives after being exposed to deep discounts in the past. In contrast, the search mechanism would have not predicted heightened promotion sensitivity if deals were negatively intertemporal correlated, as in this case treated consumers should have been less willing to search for deals, during the measurement phase, than their control counterparts. We use the estimated transition matrix of discounts to test wheter our model can replicate the observed treatment effect.

Heightened Price Sensitivity. In order to test whether our search model can rationalize heightened price sensitivity, we compare the purchase probabilities of promoted goods for consumers who faced deep discounts (treated consumers) with those who faced shallow discounts (control consumers). The model can rationalize our experimental evidence if treated consumers are more likely to buy promoted products than their control counterparts, despite

[35]Our approach is consistent with most of the empirical scanner panel literature. In contrast, imposing a Bayesian equilibrium requirement would entail having consumers form beliefs based on firm-side fundamentals. Such equilibrium outcomes are much more challenging to investigate, as they require the researcher to 1) impose a rationale on the distributions of random variables affecting firms' willingness to offer promotions, and 2) later check potential promotional deviations based on each set of potential consumer beliefs.

currently facing the same discounts.

The analysis focuses on the milk category as the implementation of the experiment closely followed the experimental design as shown in Tables 4-6. Within this category, we model the actions of the three market share leaders in the non-fat milk sub-category and assign purchases of all other types of regular (i.e. non-flavored) milk to the outside option. This sub-market choice is motivated by a number of factors. First, the three leading products belong to competing firms that hold a sizable total market share (33 percent) in the milk category and negotiate discounts independently with the retailer. Second, by including purchases of additional milk products (i.e. non-fat alternatives, 1.5 percent and whole milk offerings) our model allows for sub-market expansion effects.

The three leading brands we focused on are differentiated products. Brand 1 is the market leader and charges a price premium of 5 to 10 percent as shown in the first column of Table 15. Brand 2 is the retail chain's private label. Brand 3 is closer to a niche brand, sold at an intermediate price point with the lowest market share among the three as shown in the second column of Table 15.[36]

To compute the purchase probabilities of promoted goods of treated and control consumers, Table 14 displays the preference parameter estimates of our search model. All brand intercepts are negative, which reflects the fact that the outside option has the highest market share. In addition, the state dependence parameter implies that purchasing a product today increases the probability that the same product will be bought again in the future. The point estimates of both discount levels are positive with the high discount coefficient being larger than the shallow discount coefficient as expected. Table 15 shows that the model recovers the market shares accurately, with discrepancies being less than two percent between predicted and actual market shares.[37]

To test whether our search model can rationalize heightened price sensitivity, we compute the difference in purchase probabilities between consumers who expect deep discounts relative to consumers who expect shallow discounts, despite facing the same actual shallow discount:

$$Pr\left(Buy_j|\,\omega^D, d_j^{Shallow}=1\right) - Pr\left(Buy_j|\,\omega^S, d_j^{Shallow}=1\right) \tag{16}$$

[36]We do not include consumer visits in which no milk purchases were made in the outside option because promotional activity is unlikely to drive consumption expansion in the milk category.

[37]We considered different parameter values for search cost c and their resulting market shares to make a normalization decision as discussed in subsection 5.1. We set the search cost equal to one half.

The first term of Equation (16) corresponds to the probability of buying milk brand j given that the consumer faced deep discounts for milk in the past and hence expects a shallow discount with probability 0.556 (Column (3), Table 13). Analogously, the second term of Equation (16) corresponds to the probability of buying brand j given that the consumer faced a shallow discount for milk and hence expects a shallow promotion with probability 0.839 (Column (1), Table 13).

The results in Table 16 provide evidence that the search mechanism can rationalize our experimental findings as the differential consumer beliefs lead to a sales increase of 14.4 percent that lies in the range of the treatment effects presented in Section 3. Within our model products promoted are more likely to be bought when consumers expect there to be deep discounts. The underlying rationale is that consumers are more likely to search discounted alternatives when they expect deeper discounts, and so promoted products are searched earlier and as a result have a higher likelihood of being bought.

5.3 Implications for Competing Firms and Equilibrium Outcome

In this section we investigate the implications of our model for promotional activities of competing firms. A naive interpretation of our results suggests that heightened promotion sensitivity should increase firm profits. For example, a firm may offer a deep discount once to induce search, and then need only offer shallow discounts in the future to generate positive results on sales. This interpretation is correct, however, if firms offer discounts unilaterally. Thus, we study the equilibrium play in a game in which competing firms decide whether to offer discounts to consumers, with the understanding that these promotions affect consumer beliefs and subsequent search behavior.

Supply-Side Model. We consider a two-period model in which single-product firms compete through promotions strategically. Their action space is to sell at the regular price (i.e. no promotions), offer a shallow discount or a deep discount. Formally, in period $t \in \{1, 2\}$, firm $j \in 1, \ldots, n$ chooses a discount level $Disc_{jt} \in \{None, Shallow, Deep\}$ and faces the demand function $D_{jt}\left(Disc_{jt}, Disc_{-jt}, \Omega_t\right)$, where $Disc_{-jt}$ is a vector of discounts offered by firm j's rivals in the same period, and Ω_t is a vector of state variables that affect competition. In particular, Ω_t contains a $(n \times 1)$ vector $\widehat{\omega}_t$ that summarizes consumers' expectations about discount levels for each firm, and an additional vector with consumer state dependence accounting for previous period purchases.[38]

Each firm's objective is to maximize its expected two-period profit.[39] Equilibrium profits are given by:

$$\Pi_j^* = \pi_{j1}^* + \pi_{j2}^* \tag{17}$$

where, π_{jt}^* is defined as:

$$\pi_{jt}^* = \max_{Disc_{jt} \in \{None, Shallow, Deep\}} E\left\{ \left(p_j\left(1 - Disc_{jt}\right) - mc_j\right) \cdot D_j\left(Disc_{jt}, Disc_{-jt}^*, \Omega_t\right) \right\} \tag{18}$$

where p_j and mc_j are the firm's price and marginal cost, respectively, and $Disc_{-jt}^*$ is the optimal discount level of the j's competing firms in equilibrium.

We focus on subgame perfect equilibria, such that firms understand that their current actions influence the market conditions and the actions of their competitors in the next period. The model is solved iteratively by backward induction, i.e., solving the second period first, and then using the results to inform the first period's competition. We use the fact that the firms' actions in the first period are a sufficient statistic for the second-period state variable Ω_2. Once the second-period payoffs are calculated for each of the first-period action profiles, a deviation analysis is performed in order to identify potential equilibria of the second subgame. Tentative equilibrium-path profits are then plugged into the first period optimization problem, where all action profiles are visited again. We are able to identify all pure subgame perfect equilibria by inspecting all possible first-stage configurations. When a pure-strategy equilibrium does not exist, an exhaustive search is employed to uncover mixed strategy outcomes.[40]

Simulations. To study the impact of consumer beliefs in competition, we simulate and compare firms' actions and profits in the milk category. We compare two scenarios: A

[38] For parsimony, in the model we assume firms observe individual consumers' buying behaviors, and hence can have a precise idea of the realized state dependence in the market. In reality, firms may not have complete information on consumer behavior, and thus have to integrate over the distribution of state dependence in order to optimize their promotional offerings. Our full information assumption is analogous to the "plug-in" approach as described by Misra and Dubé (2017), albeit in a different pricing context.

[39] We do not discount second-period payoffs, given the weekly timing of promotions.

[40] Given the number of players and size of the action space, finding mixed strategy equilibria is tedious. When the analysis finds no pure strategy equilibria, a file is automatically prepared to be read by the Gambit software (`www.gambit-project.org`), which is then used to run all available algorithms to generate an exhaustive list of mixed strategy equilibria. Despite the search depth, we never found more than one mixed strategy outcome at each cost level.

scenario in which consumers are myopic (i.e. current promotions do not affect their beliefs about future ones) with a scenario in which they are sophisticated (i.e., promotions today affect their beliefs about the occurrence of promotions tomorrow).

We use the data to feed our simulations. First, we calibrate the beliefs and the discount depths following Table 13 and Table 4, respectively. In the scenario where consumers are myopic, they believe that the probability of facing a shallow discount is 0.839 in both periods for all firms. In the scenario where consumers are sophisticated, they share the same initial belief, but decrease it to 0.556 for firms who offer deep discounts in the first period. Consistent with the actual discounts in the Milk category during our experimental intervention, firms offer a regular price (no discount), a shallow promotion (4.4 percent discount) or a deep promotion (23.9 percent discount). Second, we set the regular prices p_j based on our data. Third, we further use the demand parameters estimated in subsection 5.2. Finally, the state variables in the first period, Ω_1, are set to match the firms' market share in the data.

In our simulations we explore different values for marginal costs implying a wide range of manufacturer's margins. Different manufacturer's margins can be justified by fluctuations in marginal costs (variation in production and opportunity costs) and fluctuation in prices, mainly due to promotional activities of the firm. For example, manufacturers face inventory and cost variations over time as well as expiration dates, all of which affect their opportunity costs. These forces lead economic margins to often fall above 'accounting' ones and link the price promotions to the weekly conditions of the firm. Also manufacturers' margins are sensitive to retailers' promotion policy that target specific products increasing the manufacturer's margin.

Table 17 presents the subgame perfect equilibrium action profiles at different margin levels when consumers are myopic. When consumers do not update beliefs, firms prefer to engage predominantly in shallow promotions. In the first period, firm 1 and 2 offer shallow promotions with certainty for margins below 90 percent. Only at the highest 90 percent margin level, firms 1 and 2 play a mixed strategy involving deep discounts. Instead, small firm 3 offers deep discounts with certainty at any margin above 60 percent. The reason for this reluctance to introducing deep discounts is that when consumers are myopic—and thus a deep discount does not increase their belief about future promotions and only the state dependence effect is present.—the firm looses the ability of inviting consumers to search for deals in the second period, thus decreasing the returns on offering deep discounts in the first period. The more aggresive promotional policy by the small firm 3, relative to the other larger firms, is intuitive since a low market share implies a lower position in the ranking of the sequential search and hence the largest potential gain by inviting consumers to search. Finally, notice that in the second period all firms offer shallow promotions. This is intuitive as in the second period firms can not take advantage of dynamic effects and it is not optimal to charge regular prices due to competition.

Table 18 presents the subgame perfect equilibrium action profiles at different margin levels, when consumers are sophisticated. Compared to myopic consumers, we find that firms offer deep promotions more often. In the first period firms 1 and 2 play a mixed strategy involving deep promotions when margins are 80 percent and with certainty at 90 percent margin. Small firm 3 also starts offering deep discounts with certainty at 50 percent margins, which is lower than in the previous scenario. In contrast to the case with myopic consumers, offering deep promotions in the first period can be profitable because firms can then take advantage of the state dependence effect and the heightened promotion sensitivity in the second period. Finally, notice that, just as with myopic consumers, in the second period all firms offer shallow promotions.

Table 19 shows the profit ratios between both scenarios, where a number above 100 means that the firm makes higher profits under sophisticated consumers than under myopic ones. In the 50-70 percent margin range, firm 3 is better off offering a deep discount. The rationale is that on average, firm 3 is the last one to be searched, and so is the one that can gain the most from inducing search by consumers through discounts. This has a 4 percent profitability impact on firm 1, at the 50 percent margin, but then attenuates to 0.2 percent from the 60 percent to the 80 percent margin levels. The 80 percent margin level is interesting because, under myopic consumers, firms 1 and 2 find it advantageous to offer deep discounts as well. As a result, all firms become worse off in the scenario with myopic consumers. The result is an instance of Bertrand supertraps (Cabral and Villas-Boas, 2005) by which a seeming advantage for a monopolist (offering deep discounts today in order to induce more search tomorrow) effectively decreases all competitors' profits in dynamic competition.

The results indicate that the promotional pressure sources from firms with smaller market shares first. Small firms have a higher incentive to induce search and thus affect consumer beliefs about deals in the next period. Firm 3's gains from sophisticated consumers are non-monotonic in the category's incentives to promote: at moderate margins, it benefits from being able to offer deep discounts, generating heightened promotion sensitivity in the second period. However, once bigger firms also offer deep discounts, firm 3 becomes worse off than if it faced myopic consumers. Heightened promotional sensitivity can thus benefit small firms when they are the only ones promoting by inducing additional search. However, once the whole category finds it beneficial to promote, all firms may become worse off and small firms lose the most.

6 Conclusion

By use of a large scale experiment this article has shown that managers have reasons to be worried when offering promotions. We found that deep promotions heighten customers' future promotion sensitivities. In particular, customers are 14 percent more likely to buy promoted goods after being exposed to 30 percent promotion discounts rather than 10 percent. Along the same lines, we find that the proportion of promoted goods in consumers' baskets increases in 20.5 percent, and the share of wallet also increases in 16.6 percent for treated consumers relative to their control group. The effect varies significantly across product categories, and does not seem to require a previous purchase to take place.

We show that a model of sequential search can rationalize our findings. The underlying mechanism is that our treatment changes the distribution of expected depth of the discounts, inducing treated customers to search more on promoted items than their control counterparts. Since the promotion depth was historically persistent over time in our retailer, it seems likely that treated consumers were expecting better deals among promoted items in the second half relative to their control group, explaining our main findings.

While our analysis reveals some suggestive evidence that this particular mechanism is at play in our data, our goal is not to claim that it is the only force generating the experimental results. In particular, Anderson and Simester (2004) have proposed and analyzed an exhaustive collection of additional mechanisms, some of which may be in play in our context as well. Rather, we use the experimental variation in the data to inform a well-defined mechanism of search that rationalizes our results, in order to draw strategic implications for firms.

Finally, the competition model further revealed that subsequent competition becomes fiercer after a period of deep promotions suggesting that heightened promotion sensitivity fall into the class of Bertrand supertraps explained by Cabral and Villas-Boas (2005), in which an apparent advantage for one firm can lead to lower profits (through fiercer competition) to competing firms.

Appendix

A Criteria used for Category Selection

We selected categories with the goal of providing the maximum amount of useful variation. First, we wanted to limit the influence of stockpiling behavior on the response to the promotion stimulus. If consumers respond to promotions by anticipating purchases, then post promotion dips can affect our estimates. On this basis, we excluded a few categories for which households' inventory costs were deemed to be very low (e.g., soups) and others for which consumers could keep the product in inventory for a period of time, well beyond the post-promotion period (e.g., coffee). A second related consideration for including a category was the length of the typical interpurchase time observed in the category. In particular we excluded those categories for which typical interpurchase times exceeded 5 weeks on average. Third, we only included categories that had already been promoted on a regular basis. Since our focus is on the effects of changes in promotion depth, we wanted to keep the frequency with which products were placed on promotion as constant as possible. This led us to exclude categories such as "baked goods" which were rarely, if ever, placed on promotion. Fourth, we included categories that were purchased across different demographic segments (i.e., heterogeneous in terms of socioeconomic groups and ages). By imposing this requirement, we wanted to ensure that the same categories would be relevant across all stores included in the experimental design. Fifth, we chose categories in which consumers were unlikely to use the presence of a promotion as an input in their assessment of a product's quality. It is possible that the presence of a promotion in certain categories (e.g., fresh produce) can be interpreted as a negative quality signal, e.g., the product is about to expire or does not sell well, and the promotion is seen as an attempt to sell it rapidly. Sixth, we chose categories with different degrees of brand loyalty, e.g., soft drinks are well-known for having a few star brands with very loyal consumers, whereas Milk exhibits more generic products, likely to be considered close substitutes by more consumers. Other considerations that played a role in our choice of categories were avoiding categories in which stockouts were known to occur more frequently and avoiding categories with a small number of brands.

B Model Free Analysis

It is worth considering model-free effects of the intervention, even if they do not include the efficiency from additional controls provided by regression analysis. Figure 6 depicts category-specific sales data over time. The x-axis depicts normalized time. Specifically, time 0 is the week in which a given product is sold on promotion for the first time, during the intervention phase.[41] At time 0, a given sku is sold in promotion at the 30 percent and 10 percent discount levels in treated and control stores, respectively. From then on, all promotions are shallow (10 percent). In each graph, sales of promoted items in control stores are given by black bars. The white bars depict sales of the same items sold in promotion in treated stores. Finally, the dark grey bars depict the sales of the products in control stores in non-promotional weeks, and the light grey bars depict sales of the same products in treated stores at regular prices.

The panel for the Bread category reveals that, in this case, the retailer followed our initially proposed schedule: All items are promoted in normalized weeks 0 and 5, and a clear sales spike is observable in week 0, due to the promotion differential. The panel for the Beer category shows a different promotional timing, set in advance with the retailer due to logistical constraints already discussed. In this case, the same item could be on promotion in one pair of treated/control stores, while not being in promotion on another exactly at the same time. As a result we see that, in the Beer category, items started being promoted at the 10 percent level as early as 2 weeks after the initial intervention.

Investigating the Bread and Milk categories first (i.e., the ones in which the retailer applied our initially proposed schedule across all stores), there is apparently little change in responses to subsequent promotions in the Bread category: the difference in sales across stores in week 5 is quite similar to other differences in remaining weeks, in particular to week 6. In contrast, the Milk category shows higher sales in treated than control stores in week 5, an inversion of the typical sales patterns in neighboring weeks. This happens despite the fact that both control and treated consumers face the same 10 percent discount across the same set of items.[42] The results from the Milk category suggest the intervention in normalized week 0 has led to heightened promotional sensitivity, rather than deal addiction.

[41] This time normalization generates a triangular sales pattern over time, a mechanical consequence of different sku's being promoted in different weeks.

[42] Table 5 reveals a 1 cent difference in promotional prices in favor of treated stores, which on its own is unlikely to account for the difference in behaviors.

The different schedules implemented in different categories make the comparison less clear. For example, the Cold Cuts category shows higher sales in treated than control stores for promoted items vs. unpromoted items for weeks 2, 3, 4 and 5. However, the difference in sales at regular prices between treated and control stores is higher in weeks 6 and 7, when compared to the same difference of items sold at promotional levels. Table 20 summarizes the signs of sales differences, in cases where an overlap of promotional selling conditions takes place. A rough inspection reveals more positive than negative signs, suggesting that the main effect of our intervention may be of heightened promotional sensitivity. This is of course a rough measurement: a visual inspection of sales measures does not provide a precise estimate of magnitudes, nor does it reveal information about the statistical significance of the potential results.

C Robustness

C.1 Econometric Specifications

The linear econometric specifications of our main analyses have not taken into account the fact that the dependent variables are discontinuous and/or bounded. Table 21 summarizes the main findings when the nature of the dependent variable is explicitly taken into account. In particular, we use a probit regression to estimate the treatment effects on the probability of consumers buying a promoted sku, a zero-inflated Poisson regression for quantity dependent variables, a $1 + \log(\cdot)$ transformation to expenditure variables and finally a double truncated Tobit regression for percentage dependent variables (both of which show probability masses at values 0 and 1).[43] The significance results of the treatment effects found previously carry over to the new regressions. The new treatment p-values are not calculated via the wild-bootstrap methodology. The p-values of the first three columns, in particular, fall below the comparable p-values of row (b) of Table 10.

[43]As a result, the coefficients have method-specific interpretations, and should not be directly compared with the ones found in the main analysis.

C.2 Placebo Test

In this section we introduce a placebo test designed to assess whether our experimental intervention is responsible for the differences in consumer behavior across treated and control pools. We focus the analysis on customers who did not visit the supermarket, or alternatively, did not use their loyalty cards during the first half of the experiment. Since these customers are less likely to have been exposed to the differential treatment conditions, we expect to find lower magnitudes and statistical significance of treatment effects.

The results of the analysis are presented in Table 22. No significant treatment effects are found across measures. Moreover, compared to the results of the main analysis, all estimated treatment effects fall closer to zero than before, which is in line with the interpretation that our experiment played less of a role on the behavior of these customers. Unfortunately, it is impossible to rule out that these customers may have been exposed to our intervention in the store (without using their loyalty cards).

D Search Model

D.1 Proposition: Logistic Uncertainty

The logistic p.d.f. and c.d.f. with unit scale parameter are given by $f(x) = \frac{e^{-(x-\mu)}}{\left(1+e^{-(x-\mu)}\right)^2}$ and $F(x) = \frac{1}{1+e^{-(x-\mu)}}$ respectively. We seek the solution to equation

$$z = -c + \int_z^\infty x dF_j(x) + F_j(z) z$$

with respect to z. Plugging in the expressions above yields

$$z = -c + \int_z^\infty \frac{xe^{-(x-\mu)}}{\left(1+e^{-(x-\mu)}\right)^2} dx + \frac{z}{1+e^{-(z-\mu)}}. \tag{19}$$

Integration by parts yields

$$\int_z^\infty \frac{xe^{-(x-\mu)}}{\left(1+e^{-(x-\mu)}\right)^2} dx = \log\left(e^z + e^\mu\right) + \frac{z}{1+e^{z-\mu}} - z \tag{20}$$

and the reservation value equation (19) becomes

$$
\begin{aligned}
& z=-c+\log\left(e^{z}+e^{\mu}\right)+\underbrace{\frac{z}{1+e^{z-\mu}}-z+\frac{z}{1+e^{-(z-\mu)}}}_{=0} \\
\Leftrightarrow\quad & z=-c+\log\left(e^{z}+e^{\mu}\right) \\
\Rightarrow\quad & z^{*}=\log\left(\frac{e^{\mu}}{e^{c}-1}\right)=\mu-\log\left(e^{c}-1\right)
\end{aligned}
$$

and the solution is unique for $\mu, c \in \mathbb{R}$.

D.2 Theorem: Contraction Mapping

Let v be a random variable with continuous p.d.f. and c.d.f. $f(\cdot)$, $F(\cdot)$ respectively. The indifference condition is given by

$$
\begin{aligned}
z^{*} &= -c+Pr\left(v \geq z^{*}\right) E\left[v \,|v \geq z^{*}\right]+Pr\left(v<z^{*}\right) z^{*} \\
&= -c+\int_{z^{*}}^{\infty} v f(v)\, dv+z^{*} F\left(z^{*}\right)
\end{aligned}
$$

Define $\Gamma(z)=-c+\int_{z}^{\infty} v f(v)\, dv+z F(z)$. Under standard continuity assumptions, $\Gamma(z)$ is a contraction mapping if $\Gamma'(z) \in [0,1)$. In our case,

$$
\begin{aligned}
\Gamma'(z) &= \frac{d}{dz}\left(-c+\int_{z}^{\infty} v f(v)\, dv+z F(z)\right) \\
&= 0-z f(z)+z f(z)+F(z) \\
&= F(z)
\end{aligned}
$$

which is bounded between zero and one. The use of the Leibniz integral rule implies integration must be interchangeable with differentiation. For differentiable c.d.f.'s, we only require that $zf(v)$ be continuous and a have a first derivative in z, which follows trivially. Hence, the contraction mapping applies for a large class of differentiable distributions, as long as $\int_{z_n}^{\infty} x f(x)\, dx$ is finite $\forall z_n \in \mathbb{R}$, which is also implied by the original Weitzman (1979) model. So, the theorem applies without loss of generality.

Using the proposition above, it is easy to show that in the case of mixture of logistics given by equation (13), the reservation value can be found through contraction

$$\Gamma(z) = -c + \omega_{ijt}^{H} \log\left(e^{z} + e^{v_{ijt}+\gamma^{D}}\right) + \left(1 - \omega_{ijt}^{H}\right) \log\left(e^{z} + e^{v_{ijt}+\gamma^{S}}\right)$$

where ω_{ijt}^{H} is consumer i's belief associated with finding a deep discount for a given promotional history $H \in \{S, D\}$.

D.3 Likelihood and Estimation

We now characterize the likelihood of an alternative being chosen, which involves adding over search sequences. First we rank the inside alternatives by their reservation values such that $z_1 > z_2 > ... > z_n$, where n is equal to the number of inside alternatives in the choice set. We depict the potential search paths consistent with a choice of alternative j in the diagram of Figure 1, where searching an additional option corresponds to a lateral movement, and a downward one depicts the purchase of alternative j. For a consumer to be willing to search option j with reservation value z_j, she must have inspected options with higher reservation values before and have found that it was worthwhile searching option j nonetheless. The reason is that options are ordered by their reservation values, and so if a consumer did not search option $j-1$ then she prefers not to search option j either. The sequence of events leading the consumer to arrive to node j is given by

$$\begin{aligned} z_1 > v_0 \wedge z_2 > \max\{v_0, v_1\} \wedge z_3 > \max\{v_0, v_1, v_2\} \wedge ... \wedge z_j > \max\{v_0..v_{j-1}\} \\ = z_j > \max\{v_0..v_{j-1}\} \end{aligned} \tag{21}$$

The identity above can be shown by induction. If a consumer searched option 2 for example, then $z_2 > \max\{v_0, v_1\}$. This implies the consumer also searched option 1 because

$$z_2 > \max\{v_0, v_1\} \Rightarrow z_1 > v_0$$

since $z_2 < z_1$. For the consumer to prefer option j to the options searched before, we require $v_j > \max\{v_0..v_{j-1}\}$, and so a consumer searches alternative j and considers it the best option up to that stage if and only if

$$Reach\, Node_j : \min\{z_j, v_j\} > \max\{v_0..v_{j-1}\} \tag{22}$$

Conditional on searching option j and preferring it up to that stage, many subsequent search paths can lead to a final choice of j, all of which are represented in the figure above. For example, the consumer may choose alternative j without searching any further, or do so after searching option $j+1$, options $j+1$ and $j+2$, etc. Let $Buy_{j|k}$ be each of such *subsequent* paths, where j is the chosen product and $k \geq j$ is the last product searched by the consumer. Then, the probability of choosing option j, which informs our likelihood function, is equal to

$$\begin{aligned} Pr\left(Choose_j\right) &= Pr\left\{Reach\,Node_j \wedge \left(Buy_j \middle|\, Reach\,Node_j\right)\right\} \\ &= Pr\left(\min\left\{z_j, v_j\right\} > \max\left\{v_0..v_{j-1}\right\} \wedge \left(\bigvee_{k=j}^{n} Buy_{j|k}\right)\right) \end{aligned} \tag{23}$$

We now characterize each of the paths, where movements referred to as 'down' and 'right' are related to the ones in the Figure 1:

$$\begin{aligned} Buy_{j|j} &= \underbrace{v_j > z_{j+1}}_{Path\,Down_j} \\ Buy_{j|j+1} &= \underbrace{\left(\sim Path\,Down_j\right) \wedge v_j > v_{j+1}}_{Path\,Right_j} \wedge \underbrace{v_j > z_{j+2}}_{Path\,Down_{j+1}} \\ Buy_{j|j+2} &= Path\,Right_j \wedge \underbrace{\left(\sim Path\,Down_{j+1}\right) \wedge v_j > v_{j+2}}_{Path\,Right_{j+1}} \wedge \underbrace{v_j > z_{j+3}}_{Path\,Down_{j+2}} \\ &\vdots \\ Buy_{j|k} &= \begin{cases} \left(\bigwedge_{l=j}^{k-1} Path\,Right_l\right) \wedge Path\,Down_k, & j \leq k < n \\ \left(\bigwedge_{l=j}^{k-1} Path\,Right_l\right), & j \leq k = n \end{cases} \end{aligned}$$

We have characterized the likelihood function. It remains to maximize it with respect to parameters, conditional on the data. Because utilities are probabilistic, we use simulation to generate v's and construct the likelihood. Moreover, the need to investigate multiple search paths led us to employ 10,000 draws per choice-alternative.

In order to account for heterogeneity in search sequences, we add a noise parameter $\eta \sim N\left(0,1\right)$ to the reservation values. For example, in some circumstances consumers may not include some products in their consideration sets, which is equivalent to those products featuring very low reservation values. This assumption also provides the demand function with smoothness for the purposes of the counterfactual analysis.

An additional difficulty with 'accept/reject choice simulation' is that small changes in parameters do not affect simulated outcomes, even for large sets of draws.[44] Moreover, the log-likelihood function exhibits saddle points that make finding the global maximum a challenging task.

We implement a patterned grid search across a wide range of parameter values, and ensure that the bounds sets for the parameters were never achieved during the estimation procedure. Calculation of the standard errors required additional smoothing. For this purpose, following McFadden (1989) we smoothed out the likelihood function by use of a kernel function, which in our case is analogous to adding a low-variance extreme-value noise to each v and z component.[45] For illustration purposes, suppose we observe option $n-1$ being chosen. The probability of this choice is

$$\begin{aligned}
Pr\left(Choose_{n-1}\right) &= \\
&= Pr\left\{\min\left\{z_{n-1}, v_{n-1}\right\} > \max\left\{v_0..v_{n-2}\right\} \wedge \left(\bigvee_{k=n-1}^{n} Buy_{n-1|k}\right)\right\} \\
&= Pr\left\{\min\left\{z_{n-1}, v_{n-1}\right\} > \max\left\{v_0..v_{n-2}\right\} \wedge \left(v_{n-1} > z_n \vee \left(v_{n-1} < z_n \wedge v_{n-1} > v_n\right)\right)\right\}
\end{aligned}$$

Given a parameter guess, we generate R sets of simulations of v's. For each set r, we calculate

$$p^r\left(Choose_{n-1}\right) = K\left(\min\left\{z_{n-1}, v^r_{n-1}\right\} - \max\left\{v^r_0..v^r_{n-2}\right\}\right).\left[K\left(v^r_{n-1} - z_n\right) + K\left(z_n - v^r_{n-1}\right).K\left(v^r_{n-1} - v^r_n\right)\right.$$

where

$$K\left(x\right) = \frac{1}{1 + \exp\left(-\frac{x}{\sigma}\right)}$$

is the logistic kernel with smoothing parameter $\sigma = 0.001$.[46] We used the smoothing parameter to calculate standard errors. During estimation, we used $K\left(x\right) = 1\left(x > 0\right)$ instead, because the grid search algorithm does not require smoothing out the objective function.

Finally, we average across simulation results to calculate the choice probability, i.e.

$$Pr\left(Choose_{n-1}\right) \simeq \frac{1}{R}\sum_{r=1}^{R} p^r\left(Choose_{n-1}\right).$$

[44]See Train (2009) (Sec. 5.6.2) for a careful exposition of this issue.
[45]See Honka and Chintagunta (2017) for an application within the search framework.
[46]Note that the reservation values do not require simulation because they do not depend on idiosyncratic factors.

McFadden (1989) characterizes the estimator above as well as its consistency in detail.

References

Anderson, E. T., and D. I. Simester (2004): "Long-Run Effects of Promotion Depth on New Versus Established Customers: Three Field Studies," *Marketing Science*, 23(1), 4–20.

——— (2010): "Price Stickiness and Customer Antagonism," *The Quarterly Journal of Economics*, 125(2), 729–765.

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How Much Should We Trust Differences-In-Differences Estimates?," *The Quarterly Journal of Economics*, 119(1), 249–275.

Blattberg, R. C., and S. A. Neslin (1993): "Sales promotion models," *Handbooks in Operations Research and Management Science*, 5, 553–609.

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.

Brynjolfsson, E., and M. D. Smith (2000): "Frictionless commerce? A comparison of Internet and conventional retailers," *Management science*, 46(4), 563–585.

Cabral, L. M., and M. Villas-Boas (2005): "Bertrand supertraps," *Management Science*, 51(4), 599–613.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-based improvements for inference with clustered errors," *Review of Economics and Statistics*, 90(3), 414–427.

Cameron, A. C., and D. L. Miller (2015): "A practitioner's guide to cluster-robust inference," *Journal of Human Resources*, 50, 317–372.

Chan, T., C. Narasimhan, and Q. Zhang (2008): "Decomposing promotional effects with a dynamic structural model of flexible consumption," *Journal of Marketing Research*, 45(4), 487–498.

Chen, Y., and S. Yao (2017): "Sequential Search with Refinement: Model and Application with Click-Stream Data," *Management Science*, 0(0), null.

Clay, K., R. Krishnan, and E. Wolff (2001): "Prices and price dispersion on the web: evidence from the online book industry," *The Journal of Industrial Economics*, 49(4), 521–539.

Dube, J.-P., G. J. Hitsch, and P. E. Rossi (2010): "State dependence and alternative explanations for consumer inertia," *The RAND Journal of Economics*, 41(3), 417–445.

Erdem, T., S. Imai, and M. P. Keane (2003): "Brand and quantity choice dynamics under price uncertainty," *Quantitative Marketing and Economics*, 1(1), 5–64.

Erdem, T., M. P. Keane, and B. Sun (2008): "A dynamic model of brand choice when price and advertising signal product quality," *Marketing Science*, 27(6), 1111–1125.

Freimer, M., and D. Horsky (2008): "Try it, you will like it - Does consumer learning lead to competitive price promotions?," *Marketing Science*, 27(5), 796–810.

Gedenk, K., S. A. Neslin, and K. L. Ailawadi (2010): "Sales Promotion," in *Retailing in the 21st Century: Current and Future Trends*. Springer-Verlag Eds, Berlin Heidelberg.

Goolsbee, A. (2001): "Competition in the computer industry: Online versus retail," *The Journal of Industrial Economics*, 49(4), 487–499.

Guadagni, P. M., and J. D. Little (1983): "A logit model of brand choice calibrated on scanner data," *Marketing science*, 2(3), 203–238.

Hauser, J. R., and B. Wernerfelt (1990): "An evaluation cost model of consideration sets," *Journal of consumer research*, 16(4), 393–408.

Honka, E. (2014): "Quantifying search and switching costs in the US auto insurance industry," *The RAND Journal of Economics*, 45(4), 847–884.

Honka, E., and P. Chintagunta (2017): "Simultaneous or Sequential? Search Strategies in the U.S. Auto Insurance Industry," *Marketing Science*, 36(1), 21–42.

Imbens, G. (2011): "Experimental Design for Unit and Cluster Randomized Trials," *3ie, International Initiative for Impact Evaluation.*

Imbens, G., and M. Kolesar (2015): "Robust Standard Errors in Small Samples: Some Practical Advice," *Working Paper.*

Imbens, G. W., and D. B. Rubin (2015): *Causal inference in statistics, social, and biomedical sciences.* Cambridge University Press.

Jedidi, K., C. F. Mela, and S. Gupta (1999): "Managing advertising and promotion for long-run profitability," *Marketing science*, 18(1), 1–22.

Ke, T. T., Z.-J. M. Shen, and J. M. Villas-Boas (2016): "Search for information on multiple products," *Management Science*, 62(12), 3576–3603.

Kim, J. B., P. Albuquerque, and B. J. Bronnenberg (2010): "Online demand under limited consumer search," *Marketing science*, 29(6), 1001–1023.

Kim, J. B., P. Albuquerque, and B. J. Bronnenberg (2017): "The Probit Choice Model Under Sequential Search with an Application to Online Retailing," *Management Science*, 0(0), null.

Lal, R., and C. Matutes (1994): "Retail pricing and advertising strategies," *Journal of Business*, pp. 345–370.

McFadden, D. (1989): "A method of simulated moments for estimation of discrete response models without numerical integration," *Econometrica: Journal of the Econometric Society*, pp. 995–1026.

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*, pp. 248–261.

Mela, C. F., K. Jedidi, and D. Bowman (1998): "The long-term impact of promotions on consumer stockpiling behavior," *Journal of Marketing Research*, pp. 250–262.

Misra, S., and J.-P. Dubé (2017): "Scalable Price Targeting," *Working Paper*.

Neslin, S. A., and H. J. Van Heerde (2008): "Promotion Dynamics," *Foundations and Trends in Marketing*, 3(4).

Raju, J. S. (1992): "The effect of price promotions on variability in product category sales," *Marketing Science*, 11(3), 207–220.

Rubin, D. B. (1973): "Matching to remove bias in observational studies," *Biometrics*, pp. 159–183.

——— (1979): "Using multivariate matched sampling and regression adjustment to control bias in observational studies," *Journal of the American Statistical Association*, 74(366a), 318–328.

Seiler, S. (2013): "The impact of search costs on consumer behavior: A dynamic approach," *Quantitative Marketing and Economics*, 11(2), 155–203.

Shugan, S. M. (1980): "The cost of thinking," *Journal of consumer Research*, 7(2), 99–111.

Train, K. E. (2009): *Discrete choice methods with simulation*. Cambridge university press.

Tuchman, A., H. Nair, and P. M. Gardete (2017): "Complementarities in consumption and the consumer demand for advertising," *Working Paper*.

Urbany, J. E., P. R. Dickson, and R. Kalapurakal (1996): "Price search in the retail grocery market," *The Journal of Marketing*, pp. 91–104.

Villas-Boas, S. B., and J. M. Villas-Boas (2008): "Learning, forgetting, and sales," *Management Science*, 54(11), 1951–1960.

Vuong, Q. H. (1989): "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica: Journal of the Econometric Society*, pp. 307–333.

Warner, E. J., and R. B. Barsky (1995): "The timing and magnitude of retail store markdowns: evidence from weekends and holidays," *The Quarterly Journal of Economics*, 110(2), 321–352.

Weitzman, M. L. (1979): "Optimal search for the best alternative," *Econometrica: Journal of the Econometric Society*, pp. 641–654.

Zubizarreta, J. R. (2012): "Using mixed integer programming for matching in an observational study of kidney failure after surgery," *Journal of the American Statistical Association*, 107(500), 1360–1371.

Tables

Table 2: Descriptive Statistics by Week-Store

	Mean	Std. Dev.	Minimum	Maximum
Number of Visits	14,897.4	6,736.6	3,794	29,372
Total Sales (USD)	$555,043.0	$394,387.5	$148,061.4	$1,697,815.5
No. of Items Sold	303,521.2	202,001.1	71,279	823,675
- Promoted	39,337.9	24,445.6	6,881	101,604
- Unpromoted	264,183.3	178,397.5	64,398	734,648
Average Basket Size (items)	18.9	5.1	11.0	29.6
Mkt. Share of Promoted Goods	13.1%	1.9%	9.2%	18.2%
Number of Weeks	10			
Number of Stores	12			

Table 3: Descriptive Statistics by Experimental Condition

	Treated Stores (1)	Control Stores (2)
Demographics		
Age	47.760	46.283
	(14.321)	(14.335)
Fraction Female	.653	.640
	(.476)	(.480)
Pre-experimental Expenditure (per visit)		
Total expenditure	$45.717	$44.202
	(45.745)	(45.080)
Avg. expenditure on promoted products	$3.500	$3.578
	(4.405)	(4.977)
Number of customers	60,306	60,932

Notes. Individuals in the dataset include all individuals who bought something in the first-half of the experiment for whom we observe their pre-experimental behavior. Expenditure measures converted to US dollars using the average August-October 2013 exchange rate (approximately 506 CLP/USD). Standard deviations in parenthesis.

Table 4: Price Levels in the Intervention Phase

	No Promotion			Promotion			
	Control Store (1)	Treated Store (2)	% diff	Control Store (3)	Treated Store (4)	Disc. Control	Disc. Treated
Beer	\$6.63	\$6.60	-0.5%	\$6.25	\$5.11	5.7%	22.6%
Bread	\$2.30	\$2.30	0.1%	\$2.11	\$1.74	8.1%	24.4%
Breakfast Cereal	\$3.25	\$3.24	-0.3%	\$2.84	\$2.37	12.5%	26.7%
Candy	\$3.58	\$3.58	-0.1%	\$1.83	\$1.52	**48.9%**	**57.4%**
Cheese	\$4.68	\$4.67	-0.3%	\$4.26	\$3.38	8.9%	27.5%
Cold Cuts	\$6.50	\$6.43	-1.1%	\$6.47	\$5.13	0.5%	20.3%
Cookies	\$0.83	\$0.83	0.6%	\$0.75	\$0.62	8.9%	25.6%
Cooking Oil	\$3.37	\$3.46	2.6%	\$3.21	\$2.64	4.8%	23.7%
Fruit Juice	\$1.11	\$1.11	-0.2%	\$1.06	\$0.87	4.8%	22.1%
Meats	\$8.43	\$8.16	-3.2%	\$6.81	\$5.49	**19.2%**	32.7%
Milk	\$1.16	\$1.16	0.0%	\$1.11	\$0.88	4.4%	23.9%
Pasta	\$0.93	\$0.93	-0.3%	\$0.86	\$0.68	7.3%	26.3%
Snacks	\$2.03	\$2.03	0.4%	\$1.85	\$1.51	8.7%	25.7%
Soft Drinks	\$2.02	\$2.02	-0.1%	\$1.98	\$1.60	1.9%	20.7%
Tea	\$2.67	\$2.66	-0.2%	\$2.54	\$2.09	4.7%	21.4%
Water	\$1.16	\$1.16	0.0%	\$1.08	\$0.87	6.8%	24.5%
Yogurt	\$0.34	\$0.34	-0.1%	\$0.31	\$0.26	9.2%	23.1%
Average			-0.2%			9.7%	26.6%

Notes: The table presents average prices at the category level during the first five weeks of the experimental period. All prices are expressed in US dollars (Chilean pesos converted to US dollars using the average exchange rate for the period August 2013 - October 2013, approximately 506 CLP/USD)..

Table 5: Price Levels in the Measurement Phase

	No Promotion			Promotion			
	Control Store (5)	Treated Store (6)	% diff	Control Store (7)	Treated Store (8)	Disc. Control	Disc. Treated
Beer	\$6.92	\$6.95	0.4%	\$5.95	\$5.90	14.0%	15.1%
Bread	\$2.36	\$2.36	0.2%	\$2.22	\$2.19	5.9%	7.1%
Breakfast Cereal	\$3.39	\$3.31	-2.4%	\$2.95	\$3.01	13.2%	8.9%
Candy	\$3.58	\$3.58	0.2%	\$1.93	\$1.90	**46.1%**	**47.1%**
Cheese	\$4.67	\$4.66	-0.2%	\$2.62	\$2.55	**43.8%**	**45.3%**
Cold Cuts	\$6.82	\$6.69	-1.9%	\$6.08	\$6.05	10.8%	9.5%
Cookies	\$0.84	\$0.83	-0.8%	\$0.82	\$0.83	1.7%	0.5%
Cooking Oil	\$3.60	\$3.62	0.3%	\$3.88	\$3.94	**-7.6%**	**-8.9%**
Fruit Juice	\$1.18	\$1.17	-0.2%	\$1.02	\$1.02	12.8%	13.2%
Meats	\$8.56	\$8.22	-3.9%	\$8.24	\$7.97	3.7%	3.1%
Milk	\$1.21	\$1.22	0.2%	\$1.14	\$1.13	6.1%	7.0%
Pasta	\$0.94	\$0.94	-0.2%	\$0.88	\$0.87	6.2%	6.7%
Snacks	\$2.06	\$2.05	-0.3%	\$1.88	\$1.82	8.7%	11.2%
Soft Drinks	\$2.05	\$2.04	-0.2%	\$1.91	\$1.87	6.5%	8.1%
Tea	\$2.61	\$2.60	-0.2%	\$2.54	\$2.50	2.6%	3.7%
Water	\$1.18	\$1.17	-0.5%	\$1.16	\$1.15	1.6%	1.8%
Yogurt	\$0.33	\$0.33	-0.5%	\$0.30	\$0.29	11.3%	11.2%
Average			-0.6%			11.0%	11.2%

Notes: The table presents average prices at the category level during the last five weeks of the experimental period (weeks 6-10). All prices are expressed in US dollars (Chilean pesos converted to US dollars using the average exchange rate for the period August 2013 - October 2013, approximately 506 CLP/USD).

Table 6: Price and Promotion Differences across Experimental Phases

	% Change in regular prices		% Change vs. promoted prices in control store	
	Control	Treated	Control	Treated
	[(5)-(1)]/(1)	[(6)-(2)]/(2)	[(7)-(3)]/(3)	[(8)-(3)]/(3)
Beer	4.4%	5.3%	-4.8%	-5.7%
Bread	2.5%	2.6%	4.9%	3.8%
Breakfast Cereal	4.5%	2.3%	3.7%	6.1%
Candy	-0.2%	0.1%	5.2%	3.5%
Cheese	-0.2%	-0.1%	**-38.5%**	**-40.2%**
Cold Cuts	5.0%	4.1%	-6.0%	-6.4%
Cookies	1.5%	0.1%	9.4%	9.9%
Cooking Oil	6.9%	4.5%	**20.8%**	**22.7%**
Fruit Juice	5.7%	5.7%	-3.3%	-3.9%
Meats	1.5%	0.8%	**20.9%**	**17.0%**
Milk	4.9%	5.2%	3.1%	2.2%
Pasta	0.8%	0.9%	2.1%	1.2%
Snacks	1.5%	0.8%	1.4%	-1.6%
Soft Drinks	1.3%	1.2%	-3.4%	-5.3%
Tea	-2.3%	-2.3%	-0.1%	-1.5%
Water	1.9%	1.4%	7.6%	6.9%
Yogurt	-1.0%	-1.4%	-3.2%	-3.6%
Average	2.3%	1.8%	1.2%	0.3%

Notes: The table presents average price differences (in percentage terms) across experimental phases using the respective columns in Table 4 and Table 5.

Table 7: Effect of Treatment on Customer Behavior

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expenditure on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on promoted sku's	(6) No. of non-promoted sku's bought	(7) Expenditure on non-promoted sku's
Treatment	0.042	0.014	0.016	0.303	0.366	0.762	0.624
(a): OLS std. errors	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
(b): Clustered std. errors	(0.019)*	(0.029)*	(0.014)*	(0.147)	(0.088)†	(0.134)	(0.171)
(c): (b) + bootstrap-t	(0.06)†	(0.078)†	(0.038)*	(0.314)	(0.234)	(0.322)	(0.374)
Constant	0.477**	0.218**	0.192**	2.141**	2.315**	7.685**	8.939**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.025	0.002	0.005	0.045	0.000	0.087	0.094
N. Observations:	116,635						

Notes: † p≤0.10, * p≤0.05, ** p≤0.01. p-values are in parentheses. Standard errors for p-values in (b) and (c) are clustered at the store level. p-values for the treatment effect in (c) are derived from the cluster residual bootstrap-t procedure with 50,000 draws.

Table 8: Universe of Potential Pairs

Sample:	Before Matching		After Matching			Matched pairs visiting both halves		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Store Pair	Treatment	Control	Treatment	Control	Sample	Treatment	Control	Final Sample
1	32,550	10,093	5,047	5,047	10,094	1,816	1,816	3,632
2	16,829	32,141	12,562	12,562	25,124	6,032	6,032	12,064
3	10,798	31,390	10,790	10,790	21,580	4,465	4,465	8,930
4	15,401	11,420	2,044	2,044	4,088	1,379	1,379	2,758
5	12,291	13,997	1,143	1,143	2,286	773	773	1,546
Total	87,869	99,041	31,586	31,586	63,172	14,465	14,465	28,930

Notes: Individuals matched using a Mixed Integer Programming procedure (Zubizarreta (2012)).

Table 9: Pre-treatment Covariates of Control and Treated Matched Individuals

Pre-Treatment Covariate	Control	Treatment	Difference	p-value
Average Weekly Total Expenditure (USD)	$78.07	$78.43	-$0.36	(0.13)
Average Weekly Expenditure in Promoted Categories	$25.56	$25.76	-$0.2	(0.31)
Age	47.3	47.4	-0.04	(0.07)
Fraction Female	0.66	0.64	0.03*	(0.00)
Total Number of Trips in 46 weeks	26.8	27.6	-0.81*	(0.00)

Sample Size: 28,930

Table 10: Effect of Treatment on Customer Behavior - Matched Sample

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expenditure on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on promoted sku's	(6) No. of non-promoted sku's bought	(7) Expenditure on non-promoted sku's
Treatment	0.056*	$0.036^{\dagger}$	0.034*	0.358	0.287	0.398	0.095
	(0.028)	(0.054)	(0.05)	(0.182)	(0.268)	(0.316)	(0.678)
Constant	0.452**	0.201**	0.188**	2.082**	2.249**	7.617**	9.216**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.031	0.006	0.006	0.036	0.035	0.066	0.074
N. Observations:	9,026						

Notes: † $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$. p-values are in parentheses. Standard errors are clustered at the store level and p-values for the treatment effect are derived from the cluster residual bootstrap-t procedure with 50,000 draws. The analysis is based on matched customers.

Table 11: Effect of Treatment on Customer Behavior for New Purchases

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expenditure on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on promoted sku's	(6) No. of non-promoted sku's bought	(7) Expenditure on non-promoted sku's
Treatment	0.056*	0.036*	0.032*	0.319	0.231	$0.404^{\dagger}$	0.157
	(0.026)	(0.042)	(0.038)	(0.13)	(0.274)	(0.2)	(0.526)
Constant	0.423**	0.207**	0.215**	1.9**	2.071**	6.133**	7.713**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.025	0.006	0.004	0.027	0.026	0.053	0.054
N. Observations:	8,072						

Notes: † $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$. p-values are in parentheses. Standard errors are clustered at the store level and p-values for the treatment effect are derived from the cluster residual bootstrap-t procedure with 50,000 draws.

Table 12: Discount Depth Transition Probabilities of a Focal Category

		t : Shallow Discount	Deep Discount
$t-1$:	Shallow Discount	$1-\omega^S$	ω^S
	Deep Discount	$1-\omega^D$	ω^D

Table 13: Transition Matrix of Discount Depths

	(1)	(2)	(3)	(4)
	Shallow→Shallow	Shallow→Deep	Deep→Shallow	Deep→Deep
Beer	0.636	0.364	0.444	0.556
Bread	0.973	0.027	0.667	0.333
Breakfast Cereal	0.769	0.231	0.500	0.500
Cold Cuts	0.882	0.118	0.571	0.429
Cookies	0.455	0.545	0.241	0.759
Fruit Juices	0.680	0.320	0.600	0.400
Milk	0.839	0.161	0.556	0.444
Pasta	0.947	0.053	0.500	0.500
Snacks	0.563	0.438	0.292	0.708
Soft Drinks	0.947	0.053	1.000	0.000
Tea	0.879	0.121	0.571	0.429
Water	0.886	0.114	0.600	0.400
Yogurt	0.975	0.025	N/A	N/A

Table 14: Search Model Estimates

Parameter		Estimate
Alternative-specific Constants:	α_1	-1.344*
		(0.056)
	α_2	-1.529*
		(0.058)
	α_3	-1.875*
		(0.058)
State Dependence:	β	3.227*
		(0.103)
Shallow Discount:	γ^S	0.484*
		(0.070)
Deep Discount:	γ^D	0.846*
		(0.111)
N. Customers	25,366	
N. Alternatives	3 per choice occasion+outside option	

Notes: * p-value ≤ 0.01. Standard errors in parentheses.

Table 15: Market Shares - Top 3 non-fat Milk Brands

	Avg. Price (Data, USD)	Actual Mkt. Share (Data)	Predicted Mkt. Share (Model)
Brand 1	$1.10	15.08%	15.34%
Brand 2	$0.99	11.49%	11.41%
Brand 3	$1.04	6.77%	5.41%

Notes: Average prices in US dollars (Chilean pesos converted to US dollars using the average exchange rate for the period August 2013 - October 2013, approximately 506 CLP/USD). Actual market shares computed from the data. Predicted market shares obtained using our structural model.

Table 16: Counterfactual Analysis: Heightened Price Sensitivity

	$Pr\left(Buy_j\|\omega_{ic}^S\right)$	$Pr\left(Buy_j\|\omega_{ic}^D\right)$	Relative Increase
Brand 1	11.48%	13.10%	14.15%
Brand 2	10.33%	11.85%	14.70%
Brand 3	7.96%	9.12%	14.57%
Average:	9.93%	11.36%	14.40%

Notes: The table presents predicted market shares for the three top brands in the non-fat milk sub-category for the case in which firms offer shallow discounts. Predicted market shares in column 1 assume the consumer expects a shallow discount with probability Pr(Shallow | Shallow). Predicted market shares in column 2 assume the consumer expects a shallow discount with probability Pr(Shallow | Deep). Column 3 is equal to ((2)-(1)/(1)).

Table 17: Competition Counterfactuals: Myopic Consumers

- a) Subgame Perfect Equilibrium:

Period 1

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
Firm 1	S	S	S	S	S	S	$\{S : 68\%; D : 32\%\}$
Firm 2	S	S	S	S	S	S	$\{S : 64\%; D : 36\%\}$
Firm 3	S	S	S	D	D	D	D

Period 2

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
Firm 1	S	S	S	S	S	S	S
Firm 2	S	S	S	S	S	S	S
Firm 3	S	S	S	S	S	S	S

- b) Equilibrium Payoffs, per potential customer, 2 weeks (USD): $\pi^*_{j1} + \pi^*_{j2}$

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
Firm 1	\$0.11	\$0.16	\$0.20	\$0.24	\$0.28	\$0.32	\$0.36
Firm 2	\$0.08	\$0.11	\$0.14	\$0.17	\$0.20	\$0.23	\$0.26
Firm 3	\$0.05	\$0.06	\$0.08	\$0.10	\$0.12	\$0.15	\$0.16

Table 18: Competition Counterfactuals: Sophisticated Consumers

- a) Subgame Perfect Equilibrium:

Period 1

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
Firm 1	S	S	S	S	S	$\{S : 71\%; D : 29\%\}$	D
Firm 2	S	S	S	S	S	$\{S : 49\%; D : 51\%\}$	D
Firm 3	S	S	D	D	D	D	D

Period 2

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
Firm 1	S	S	S	S	S	S	S
Firm 2	S	S	S	S	S	S	S
Firm 3	S	S	S	S	S	S	S

- b) Equilibrium payoffs, per *potential* customer (assumes 1 purchase per customer at most), 2 weeks (USD): $\pi^*_{j1} + \pi^*_{j2}$

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
Firm 1	\$0.11	\$0.16	\$0.19	\$0.24	\$0.28	\$0.32	\$0.36
Firm 2	\$0.08	\$0.11	\$0.14	\$0.17	\$0.20	\$0.23	\$0.26
Firm 3	\$0.05	\$0.06	\$0.08	\$0.11	\$0.13	\$0.14	\$0.15

Table 19: Competition Counterfactual: Sophisticated Consumer Profits over Myopic Consumer Profits

Firm \ Margin	30%	40%	50%	60%	70%	80%	90%
Firm 1	100.0%	100.0%	96.0%	99.8%	99.8%	98.8%	98.7%
Firm 2	100.0%	100.0%	99.8%	99.7%	99.7%	97.7%	100.0%
Firm 3	100.0%	100.0%	101.9%	102.6%	102.6%	95.2%	92.5%

Notes: The table presents the ratio between firms' profits under the assumption that consumers are sophisticated (i.e., they do update their beliefs of future discounts based on the observed discounts in period 1) and profits under the assumption that consumers are myopic (i.e., they fail to update their beliefs of future discounts based on the observed discounts in period 1).

Table 20: Direction of Sales Differences between Promoted and Unpromoted Items

	1	2	3	4	5	6	7	8	9
Beer		$+$		$+$	$-$		$+$		
Bread									
Breakfast Cereal		$+$	$+$				$\simeq$	$-$	
Cold Cuts		$+$	$+$	$+$	$+$	$-$	$-$		
Cookies					$+$		$+$		
Fruit Juice	$+$	$-$			$+$	$+$	$\simeq$		
Milk									
Pasta		$+$					$\simeq$		
Snacks					$-$	$+$			
Soft Drinks		$+$	$+$		$-$	$+$			
Tea				$+$	$+$				
Water				$+$	$+$				
Yogurt					$-$	$\simeq$	$-$	$-$	$\simeq$

The element above sign the differences between sales of promoted and unpromoted goods across types of stores, when conditions overlap. For example, Beer is signed with a '+' on the second week, because the difference between sales of promoted sku's between treated and control stores is higher than differences of unpromoted sku's during the same week in the same stores (see also Figure 6).

Table 21: Effect of Treatment on Customer Behavior - Alternative Specifications

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expenditure on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on promoted sku's	(6) No. of unpromoted sku's bought	(7) Expenditure on unpromoted sku's
Method:	Probit	Double-Censored Tobit	Double-Censored Tobit	Zero-Inflated Poisson	Linear Reg. w/ log(1+D.V.)	Zero-Inflated Poisson	Linear Reg. w/ log(1+D.V.)
Treatment	0.144**	0.089**	0.086**	0.067	0.096*	0.074	0.001
	(0.000)	(0.001)	(0.000)	(0.448)	(0.03)	(0.106)	(0.913)
Constant	-0.123**	-0.939**	-0.102**	1.473**	0.679**	2.051**	1.904
	(0.007)	(0.018)	(0.01)	(0.000)	(0.000)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
N. Observations:	9,026						

Notes: † $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$. p-values are in parentheses. Standard errors are clustered at the store level.

Table 22: Effect of Treatment on *Placebo* Customers' Behavior

	(1) Bought promoted sku	(2) % Items bought in promotion	(3) % Expenditure on promoted sku's	(4) No. of promoted sku's bought	(5) Expenditure on promoted sku's	(6) No. of unpromoted sku's bought	(7) Expenditure on unpromoted sku's
Treatment	0.034	0.023	0.027	0.195	0.235	0.176	0.072†
	(0.324)	(0.316)	(0.756)	(0.24)	(0.21)	(0.62)	(0.26)
Constant	0.184**	0.117**	0.118**	0.539**	0.738**	4.117**	5.508**
	(0.000)	(0.001)	(0.000)	(0.047)	(0.001)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.01	0.01	0.01	0.006	0.012	0.013	0.018
N. Observations:	2,596						

Notes: † $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$. p-values are in parentheses. Standard errors are clustered at the store level and p-values for the treatment effect are derived from the cluster residual bootstrap-t procedure with 50,000 draws. Columns 4 and 5 only include treated-control pairs who have bought items during the second half of the experiment.

Figures

Figure 2: Timing of Discounts

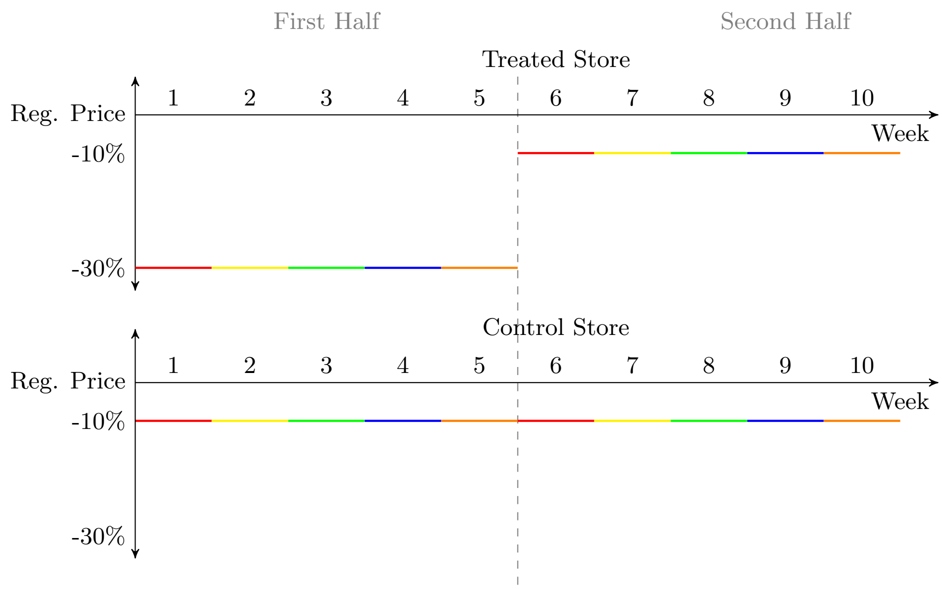


Figure 3: Pre-experimental Shopping Behavior across Experimental Conditions

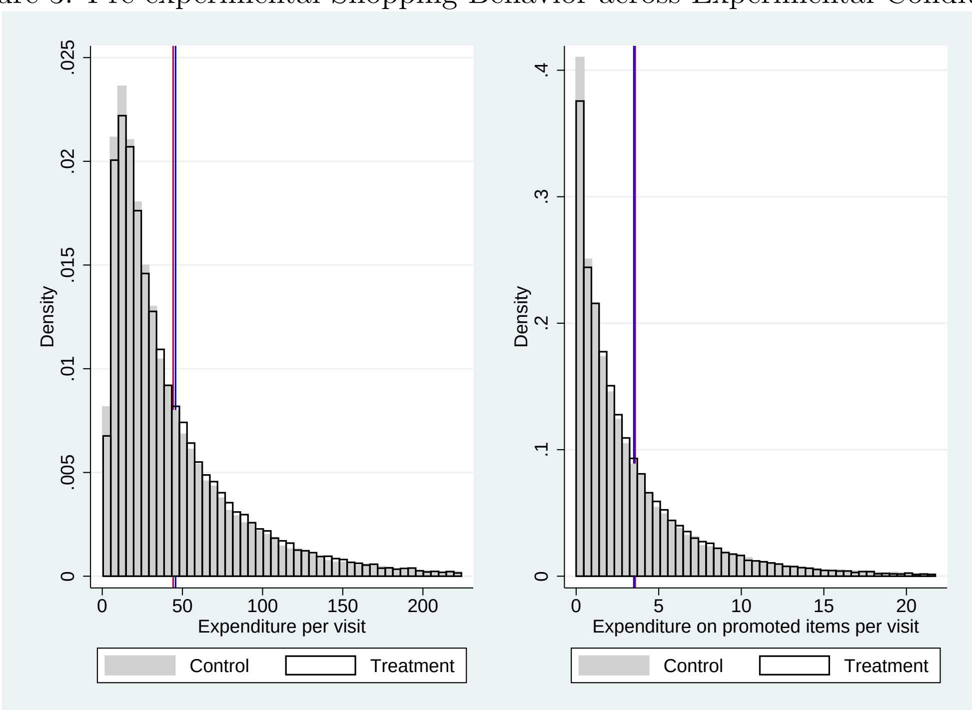
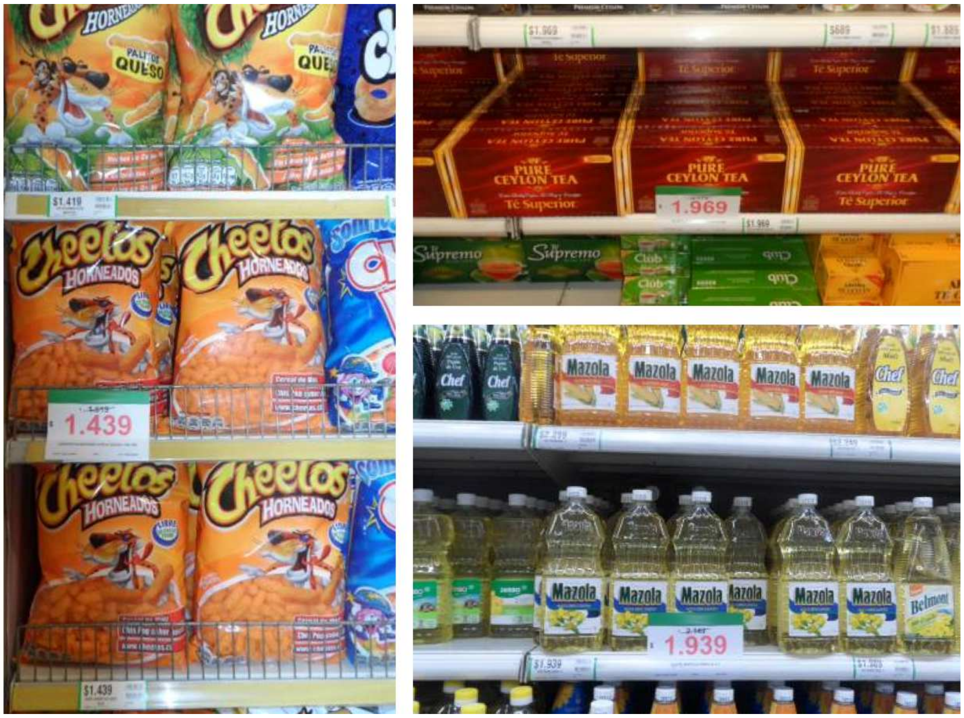
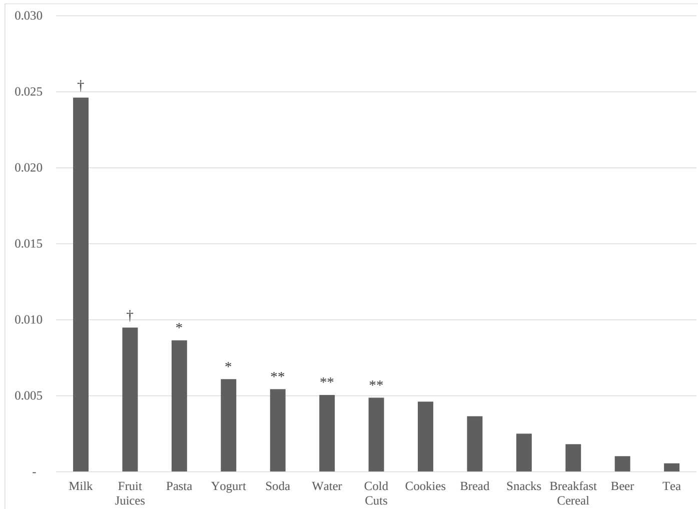


Figure 4: Promotions on Shelves



Clockwise: Promotions in Snacks, Tea and Cooking Oils categories.

Figure 5: Category-Level Treatment Effects



Notes: † p≤0.10, * p≤0.05, ** p≤0.01. p-values are in parentheses. Standard errors are clustered at the store level and p-values for the treatment effect are derived from the cluster residual bootstrap-t procedure with 10,000 draws. The analysis is based on matched customers.

Figure 6: Average Category Sales per Store Across Normalized Weeks

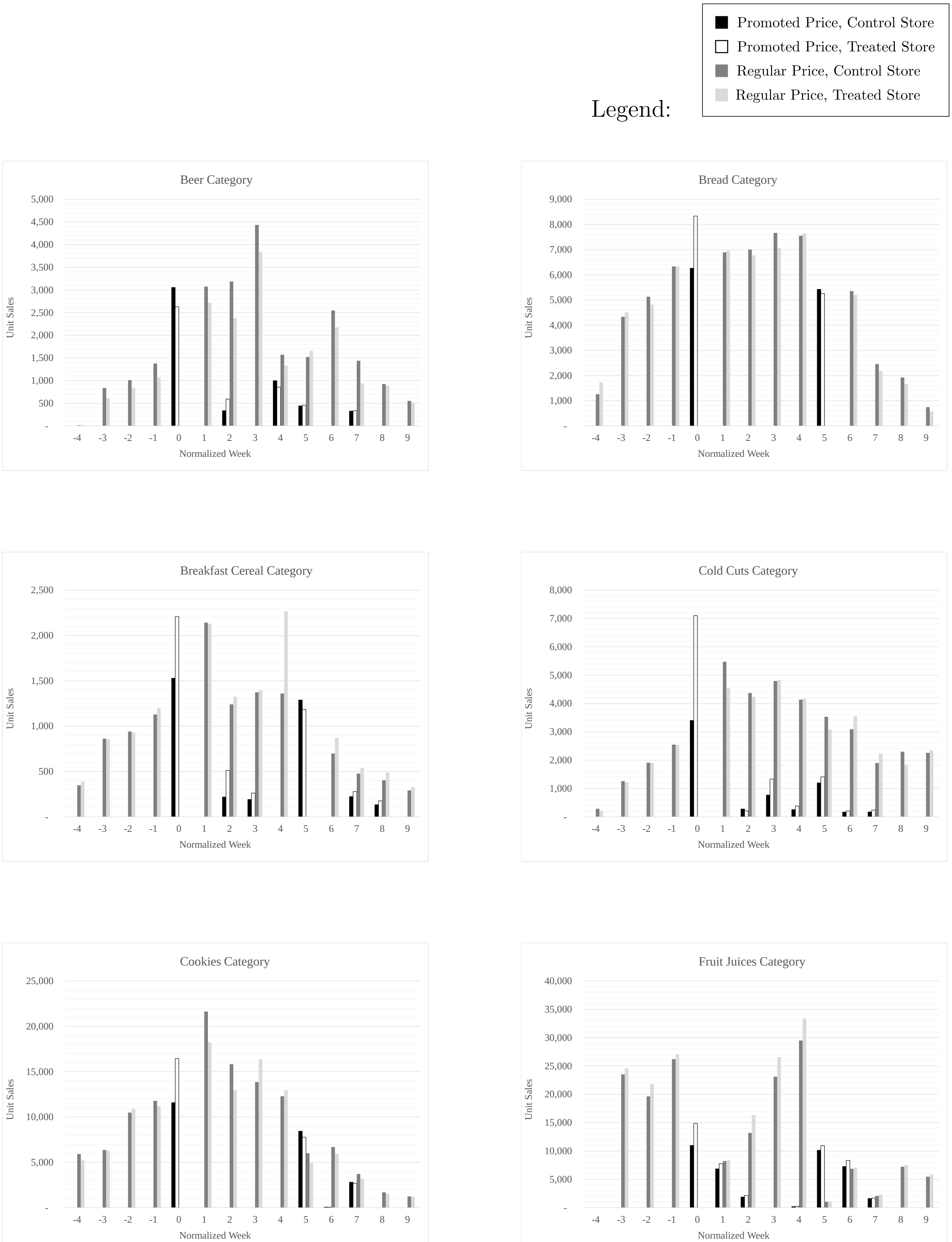


Figure 6: Average Category Sales per Store Across Normalized Weeks (contd.)

Legend:

- Promoted Price, Control Store
- Promoted Price, Treated Store
- Regular Price, Control Store
- Regular Price, Treated Store

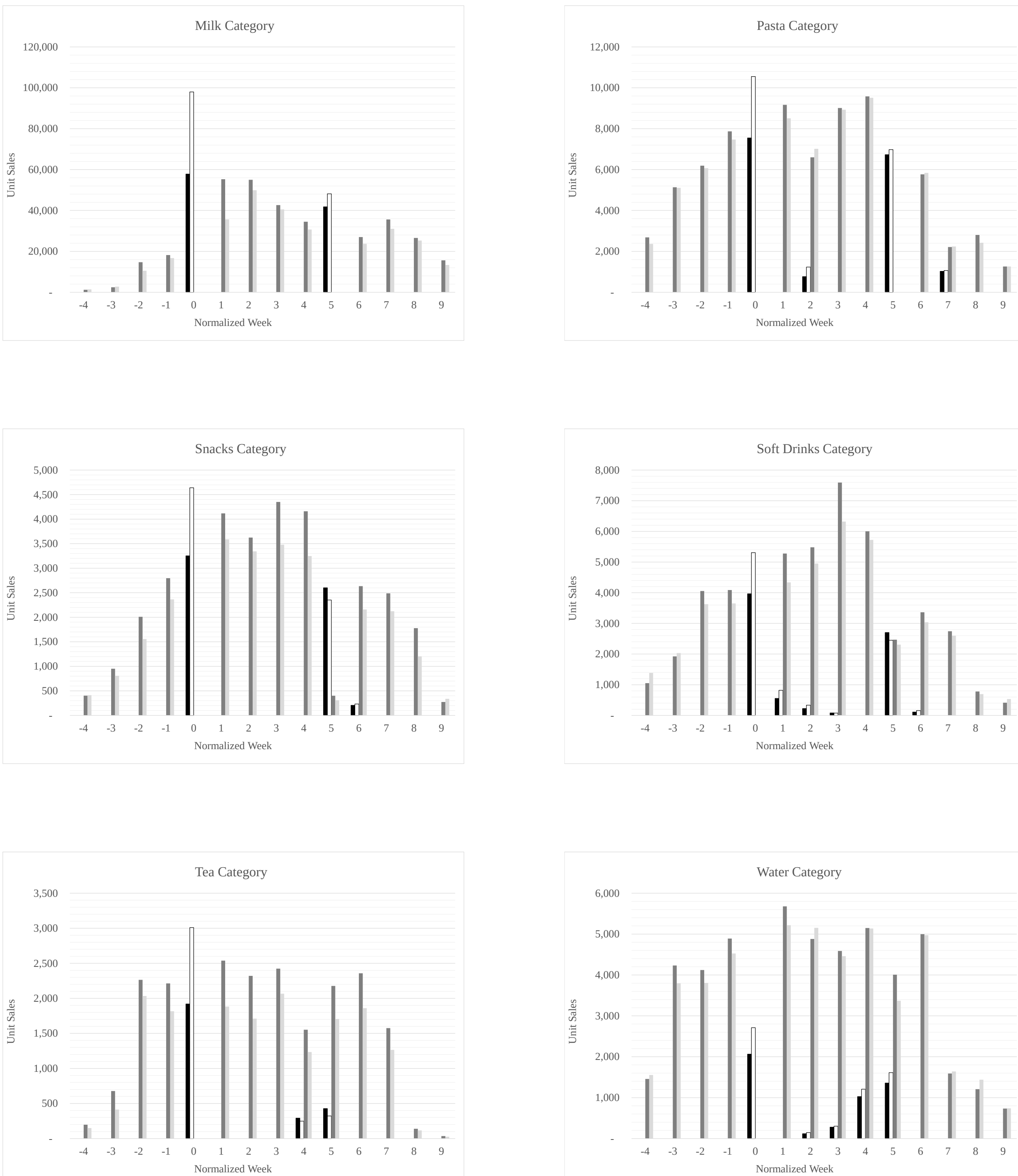


Figure 6: Average Category Sales per Store Across Normalized Weeks (contd.)

Legend:

- Promoted Price, Control Store
- Promoted Price, Treated Store
- Regular Price, Control Store
- Regular Price, Treated Store