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

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

June 2017

Abstract

This paper investigates the dynamic effects of price promotions in a retail setting through the use of a large-scale field experiment which involved varying the promotion depths of 170 products across 17 categories in 10 supermarkets of a major retailer in Chile. In the intervention phase of the experiment, customers were exposed to a promotion schedule that differed only on promotional depths: treated customers were exposed to 30% discounts, whereas control customers were exposed to 10% discounts. In the subsequent measurement phase, the promotion schedule held discount levels constant across groups. We find that treated customers were 22.5% more likely to buy promoted items than their control counterparts, despite facing the same promotional deals. The result is robust to a number of concurrent dynamic forces, including consumer stockpiling behavior and state dependence. We assess the implications to manufacturers by considering a demand-side model in which consumers search for deals, and a supply-side model in which firms compete for those consumers. We find that small manufacturers can benefit from heightened promotion sensitivity by using promotions to induce future consideration. However, when unit margins are high, heightened promotion sensitivity leads to fierce competition, making all firms worse off.

*We thank Eva Ascarza, Bryan Bollinger, JP Dubé, Przemyslaw Jeziorski, Carl Mela, Ilya Morozov, Harikesh Nair, Duncan Simester, and Miguel Villas-Boas for helpful comments, as well participants in research seminars at Pontificia Universidad Católica de Chile, Santa Clara University, UC Berkeley, Universidad Diego Portales, and at the Marketing Science Conference, SICS, the UCSD Rady Field Experimentation Conference, and the Workshop in Consumer Analytics. The outstanding research assistance by Claudio Palominos and Breno Vieira is gratefully acknowledged. Noton acknowledges financial support from the Institute for Research in Market Imperfections and Public Policy, ICM IS130002.

†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, Alice and Bob, in all identical except in the promotions they face. For example, Alice might be offered deeper discounts for a set of goods than Bob, for a period of time. According to the law of demand, Alice may be reasonably expected to demand more goods than Bob during the promotional period. In this paper we investigate the dynamic effects of promotional activities. In other words, our focus is how exposure to distinct promotional activities affect 1) Alice and Bob's reactions to subsequent promotional activities and 2) the profitability of firms competing for their business.

The answer to this question is not straightforward, and different possibilities have been proposed. For example, lay managerial theories often highlight the dangers of competing through price promotions. According to this view, a period of attractive promotional incentives may *dampen* consumers' sensitivities toward subsequent promotions, a scenario we refer to as 'deal addiction.' In this case, firms are required to ratchet incentives up over time in order to maintain buying behaviors. On the other hand, it is also possible that promotional incentives *increase* customers' promotion sensitivities. In this case, consumers exhibit *heightened promotion sensitivity*, such that they require smaller promotional incentives in order to maintain their purchasing behaviors. The implications of these scenarios are extremely broad, as they affect firms' decisions about pricing (Anderson and Simester, 2004, 2010), about promotions and advertising (Kopalle, Mela, and Marsh (1999), Erdem, Keane, and Sun, 2008), and may also affect firms' understanding of heterogeneity of consumer responses to marketing mix decisions (Chan, Narasimhan, and Zhang, 2008).

We investigate the dynamic effects of price promotions by use of a large-scale field experiment. In collaboration with a major retail chain in Chile, we exogenously varied the prices of 170 products across 17 categories in 10 supermarkets to study how the depth of present promotions affects consumers' sensitivities to subsequent promotion activities. The experiment is organized in two halves. In the first half (the intervention phase), a set of products sold in 5 treated stores received "deep" 30% discounts according to a promotion schedule. Simultaneously, the same products were sold under the same promotion schedule at "shallow" 10% discounts in 5 similar control stores. In the second half (the measurement phase), the same set of products was then attributed identical 10% discounts across all stores, and the promotion schedule was implemented once again. Our focus is on systematic consumer behavior differences occurring during the second half of the experiment, as a function of the conditions assigned during the first half.[1]

In order to assess the implications of our findings, we develop and estimate a model in which consumers search for deals. The model takes advantage of the exogenous variation generated by the experiment, and also incorporates historical data on promotional activities in order to inform beliefs over the distribution of promotions consumers expect to observe. We use the resulting structural demand function to consider the strategic interaction of firms competing over customers through promotional activities, through a supply-side model.

Our experiment reveals that consumers exposed to deep promotions exhibit subsequent heightened promotion sensitivity. In particular, treated customers were 22.5% more likely to buy promoted items during the second half of the experiment than their control counterparts, despite being exposed to the same promotions at the time. Moreover, exposure to deep promotions during the first half of the experiment also increased the proportion of promoted goods in treated consumers' baskets in 5.1% during the measurement phase.

Our findings also motivate the interpretation that offering attractive promotions today can act as an invitation to search for deals to a greater extent tomorrow. In particular, we find that promotion depths are positively correlated over time, which implies that finding a deep discount on a given week is a positive indicator of the likelihood of finding a deep discount in the following one. Under beliefs consistent with these data, the model predicts that consumers in treated stores should search and buy a higher share of promoted goods than their control counterparts. In terms of the supply-side, the counterfactual analysis reveals that small firms, whose products are the least searched, have a higher incentive to provide deep promotions in order to generate future consumer consideration. However, these firms are also the most penalized once competitors also find it valuable to offer deep promotions. In addition, we find that all firms can become worse off when consumers exhibit heightened promotion sensitivity, due to the resulting competition intensity.

This paper contributes to three research streams of the marketing literature. First, to the best of our knowledge, our paper is the first one to provide experimental evidence of the causal link between promotional activity and subsequent promotion sensitivity in a physical

[1]The marketing literature distinguishes between several price promotion instruments, including temporary price reductions (TPRs), coupons, promotion packs, rebates, among others. Our focus is on TPRs – the most frequently used type of price promotion (Gedenk, Neslin, and Ailawadi, 2010).

retail setting. Previous work has nonetheless devoted attention to this research question, by use of scanner panel data.[2] Mela, Gupta, and Lehmann (1997) initiate this line of inquiry by using 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.

The work by Anderson and Simester (2004) tackles a similar research question by randomizing prices, in the context of mail order catalogs selling durable goods. They find that new customers who are offered deeper discounts purchase more in subsequent orders relative to the control group, whereas established customers react in the opposite way by reducing their subsequent purchases. Our results complement those of Anderson and Simester (2004), since our context, focus, and experimental design feature relevant differences. In terms of the context, differences in the institutional settings (e.g., durable goods sold by catalog vs. non-durable goods sold in supermarkets) imply that the findings from one environment need not necessarily extend to the other. For example, the presence of immediate competitors in a limited shelf space, or the limited amount of consumer learning in retail settings, may lead to different forces and different results, than when compared with the context investigated by Anderson and Simester (2004).[3] Second, our main focus is also different in that we aim to analyze the implication of dynamic consumer behavior for the supply-side, and in particular for competing firms. Finally, in terms of the experimental design, our focus is on how promotion sensitivity is affected by the depth of past promotions. In contrast, the main focus of Anderson and Simester (2004) is on the subsequent purchasing behavior as a whole.[4]

We also contribute to the research on supply-side promotions by investigating the implications of heightened promotion sensitivity on promotional activities and equilibrium profits

[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]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."

of competing manufacturers.[5] Our findings suggest that small firms benefit first from offering price promotions as unit margins increase. Moreover, heightened promotion sensitivity can increase the likelihood of promotions, potentially making all firms worse off. Whereas a monopolist could potentially benefit from heightened promotion sensitivity by offering deep discounts for a short period of time followed by a series of shallow discounts, rival firms instead race to the bottom until they find themselves in a state of fierce competition. Hence, we uncover a prisoner's dilemma that rationalizes the occurrence of promotional activity, despite its damaging effects. In particular, this finding is an instance of a Bertrand supertrap, 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 competing firms.

Finally, our paper also contributes to the literature on consumer search. While precise search data is usually unavailable in retail settings, we show that normalized preference parameters of the demand model induced by search can still be estimated by integrating over consumer search paths. Recovering the fundamental demand parameters allows us to investigate a precise demand mechanism, as well as to consider the supply-side equilibrium implications to competing firms. Moreover, the demand model not only makes the demand mechanism precise, but also lends itself naturally to allow complex dynamic forces, such as the impact of price promotions on consumers' beliefs and the resulting future demand.

We provide two additional results that can be useful for the estimation of consumer search models. In particular, we build on the seminal application by Kim, Albuquerque, and Bronnenberg (2010), who estimate the sequential search model proposed by Weitzman (1979).[6] One of the key steps in estimating these models is the calculation of alternative-specific 'reservation values,' which are used to determine consumers' search sequences across products. We propose two alternative methods to the current interpolation approach. First, we show that in situations where the researcher is willing to consider agents face Logistic preference shocks, the reservation values are directly obtainable through a convenient closed-form expression. The Logistic distribution is attractive since it features an unbounded support, and is often used as an analytical approximation of the Normal distribution.

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

[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.

Second, in situations where the researcher prefers a specific uncertainty structure, we prove a contraction mapping result which implies that, in general, the reservation values are obtainable through fixed-point iteration. We show this method is applicable to a broad class of continuous distributions.

The remainder of the paper is organized as follows. The next section describes the experimental design, the data, and validates the experimental intervention. Section 3 presents the results of the field experiment. Section 4 describes the demand model, identification, estimation, and results. Section 5 presents the supply-side model and the counterfactual analyses. Section 6 concludes.

2 Field Experiment

2.1 Experimental Design

The experiment took place during the August-October period in 2013, and was conducted in partnership with a large retail firm in Chile, which has a 30% market share nationwide. The retailer's stores are organized into two retail sub-chains, which make use of different branding and perform separate marketing activities. In this paper, we report the experimental results of the intervention on the larger chain, where the depth of promotions of 170 top-selling products were manipulated across 10 stores, approximately one third of the stores held by the chain at the time.[7]

The intervention affected products in 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] 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 depths, we ran-

[7]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. The intervention did not to produce statistically significant results for stores of the smaller chain. However, all results are directionally consistent across chains and dependent variables. We believe the smaller sizes of stores in the second chain are responsible for the absence of statistically significant effects. The results of the intervention in the smaller chain are available from the authors.

[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.

domly assigned stores to one of two alternative conditions: (i) in “treated” stores, participating products featured deep discounts (30%) during the intervention phase (5 weeks), and shallow discounts (10%) during the measurement phase (5 weeks); (ii) in “control” stores, participating products featured shallow discounts (10%) during both experimental phases.[10] The discount levels of 10% and 30% are near the lowest and highest discount levels typically offered by the retailer. We offer discounts in both conditions 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 3. Each color/pattern 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% in treated stores, 10% in control stores) with promotional tags. By the end of week five, the prices of all participating sku’s had been intervened with. The same promotion schedule was then repeated during the measurement phase, this time with equal markdowns of 10% in both types of stores. As we detail later, logistical reasons led us to opt for an approximate schedule to the one depicted in Figure 3 during the second half of the experiment, in order to take the managers’ logistical constraints into account.[11]

Our experiment relies on pairwise randomization, i.e., the full sample of stores was divided into pairs according to covariate values, and each pair member was then assigned to the treatment or control group randomly (see Imbens, 2011). Out of all of the retailer’s stores, we selected store pairs along three dimensions: (i) similar consumer demographics

[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 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 to be implemented for the whole 10 weeks, 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 sales targets 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 local demand conditions. Moreover, they are constant across control/treated store pairs, as we further describe in Section 2.3, and so are unlikely to affect the validity of our estimates.

(age, gender, socioeconomic group); (ii) similar competition intensity; (iii) similar geographic location. The retailer provided us with store-level demographic information, as well as its own classification of neighborhood factors to construct the measures. Finally, the tentative store assignments were discussed with the retailer to ensure that managerial knowledge was also taken into account. This resulted in the creation of six pairs of comparable stores which, as we explain later, yielded 5 usable store pairs. After the procedure was completed, each store in each pair was assigned to an experimental condition via coin flip.

2.2 Data

The resulting scanner data provided by the retailer covers all transactions carried out by loyalty program members in the participating stores within the experimental period. The data include quantities purchased and the actual prices paid by a given customer for each sku across product 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%). We were also provided with cardholders' demographic information including their gender, age and socio-economic group classification.

The analysis focuses on purchases of experimental goods by a sample of 234,063 loyalty program cardholders that meet the following conditions: 1) bought at least one (potentially non-experimental) sku in the top 31 retailer categories during the first half of the experiment and 2) did not visit more than one of the retailer's stores during the whole experimental period. The goal of the first condition is to focus on consumers who shop relatively frequently at the retailer, and who may have been exposed to the intervention. The second condition is imposed in order to guarantee that the same consumer was not exposed to different experimental conditions. Table 1 shows the descriptive statistics of the data during the measurement phase, per store-week pair. On average, each store had about 10,000 customer visits each week, and sold roughly 70,000 sku's weekly.[12]

[12] Throughout the paper we use the symbol '$' to denote US Dollars, which were calculated by converting from the local Chilean Peso currency, based on the exchange rate of 0.0016 USD/CLP obtained from Google on May 27th, 2015.

In order to verify whether consumers exposed to different experimental conditions are comparable, the retailer made available a supplemental dataset containing individual-level behavior taking place 46 weeks before the intervention period. The dataset contains individual-level behavioral statistics for 10 weeks of loyalty program cardholders, including quantities purchased in each category.

We compare the historical baseline shopping behavior of approximately 221,323 consumers for whom we have historical data available, out of the sample of 234,063 experimental consumers. Figure 4 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 expenditures in both promoted and non-promoted categories appear to be very similar across experimental conditions, as can be gauged from the proximity between the vertical lines, blue and red, depicting per-visit 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 (age and gender), although relatively small differences exist at the behavioral level. As we explain later, these differences led us to also perform the analysis on a subsample of consumers that are matched on demographics and in behavioral measures.

2.3 Compliance with the Experimental Design

We actively engaged with the retailer in order 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 promotions they can offer, the experiment was generally well received even by managers of control stores. An executive from the retailer headquarters was named as the coordinator and supervisor of the experiment. This executive was in charge of ensuring that price lists were sent to the participating stores in a timely fashion. Finally, store managers were asked to write back to headquarters and submit photos of updated price promotions on a weekly basis, as depicted in Figure 5.

Although 12 stores were originally selected to participate, only 10 stores fully took part in the experiment. The reason was that a manufacturer noticed the effects of the intervention, which apparently benefited a major rival in a specific product category. The concerns targeted two stores, which were immediately removed from the experiment. The two stores had originally been assigned to different store pairs and different conditions. However, the fact that these pairs shared the same classification based on customer demographics and competition level allowed us, with the agreement of management, to reassign their couples to a new store pair. This is unlikely to compromise the experimental design and consistency of point estimates.

To assess compliance, we compare the average price levels that were effectively implemented against the prices defined 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 3 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 products in the non-promoted and promoted weeks are in the leftmost and rightmost columns of the table, respectively. The average price difference of non-promoted products between treatment and control stores is -0.2%, close to the ideal of the experimental design. Similarly, the average discount of promoted products was 9.7% and 26.6% in treated and control stores, respectively, relatively close to the ideal intervention of 10% and 30% discounts.

While the results related with the intervention phase were generally deemed as satisfactory, concerns related to the Candy and Meats categories emerged. Table 3 shows that the Candy category was promoted at 50% discount levels in both treated and control stores, and that the Meats category was promoted with discounts of nearly 20% in control stores. While we are not aware of the specific reasons for these deviations, we preferred to keep the analysis conservative by removing these potentially problematic categories. The reason is that, even if the phenomenon responsible for the deviations is exogenous, deviations from our intervention plan are counterproductive in answering the research question of the differential dynamic effects of offering shallow vs. deep promotions.

Table 4 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, the difference between prices across control and treated stores for the non-promoted products is less than 1%. Equally satisfactory is the fact that the average discount of promoted products was 11.2% and 11.0% in treated and control stores, respectively. Although there exists some variation on the actual promotion level offered across categories, the differences in promotion levels are relatively small across types of stores. However, a few categories exhibit unexpected patterns: the Candy and Cheese categories feature extremely high discounts, and the Cooking Oil category was sold under promotion at above average prices. This result led us to further remove the Cheese and Cooking Oil categories from the analysis.

Finally, we also compare prices across experimental halves. Table 5 shows the comparison of the average prices between the intervention and measurement phases for the 17 categories. The price comparison for the non-promoted and promoted products are in the leftmost and rightmost columns of the table, respectively. The first two columns of Table 5 reveal that regular prices increased slightly across the experimental halves, with average increases of 2.3% and 1.8% in control and treated stores, respectively. Within-category differences fall within single digits, and as a result, we see no reason to eliminate observations based on these figures. The third column, however, reveals that promoted prices in control stores changed significantly in the Cheese, Cooking Oil and Meats categories. The expected change was equal to zero, because products in control stores are supposed to be promoted at the 10% level throughout the experiment. Moreover, in the second half of the experiment, products in treated stores should also be promoted at 10%. The fourth column of Table 5 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 higher than expected differences. The levels of regular and promotional prices in the remaining categories were accepted as complying with the experimental protocol, in light of the intervention complexity.[13]

[13] All of the results of interest survive the inclusion of the problematic categories, although the significance of the treatment effects decreases. Moreover, a few confounding results arise, related to unexpected significance of a few variables with no apparent explainable pattern.

3 Results

We analyze mean differences in consumer behavior between the control and treated stores during the measurement phase (second half) of the experiment. The main model used for the analysis is given by

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

where y_i is a customer behavior measure of interest, and T_i is a dummy variable, equal to one for consumers shopping in treated stores. Parameter β_1 thus captures the change in y_i induced by exposure to 30% discounts, rather than 10%, during the first half of the experiment. 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.[14] The analysis focuses on 7 behaviors of interest during the measurement half of the experiment: whether consumers are more likely to buy a promoted item (variable 1), the proportion (quantity and expenditure) of promoted vs. non-promoted items bought (variables 2 and 3), the quantities and expenditures on promoted and non-promoted items (variables 4-7). The first 3 variables are of special interest to sign the effects of the intervention, whereas the last 4 provide insight of the underlying behaviors that generate the results.

Two notes are in order: first, given the nature of our setting, administering treatment conditions at the individual level is unfeasible. Hence, notwithstanding the magnitude of the intervention, it is important to recall that treatment conditions were assigned at the store level. As a result, we should not assume that each consumer observation in regression (1) provides an independent draw from the population. Rather, consumers of a given store may share observable and unobservable characteristics, making the data heteroscedastic at the observation level. Hence, each consumer's effective contribution to reducing standard error estimates is likely to be lower than in the i.i.d. case. We cluster standard errors at the store level in order to avoid downward biases in our standard error estimates (see Bertrand, Duflo, and Mullainathan (2004) for an in-depth discussion).[15] Second, the validity of standard error

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

[15]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 obtained through OLS and the the ones obtained through clustering at the store level.

clustering relies on the asymptotic behavior of the estimator at the cluster rather than at the individual level. Given the relatively small number of stores/clusters available for the experiment, 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 potentially induced in small samples.[16]

Table 6 summarizes the results for each of the seven measures of interest. The first three columns summarize the results of interest, whereas columns 4-7 shed light on the underlying mechanism. All point estimates in columns 1-3 are positive, suggesting consumers exhibit heightened promotion sensitivity after being exposed to deep promotions. Treated consumers were more likely to buy promoted items (column 1) than their control counterparts. They also have a higher share of promoted items in their baskets during the measurement phase of the experiment (columns 2 and 3). In terms of the underlying forces, consumers exhibit higher purchase rates of all items, independently of promotional condition (columns 4-7).

Rows (a)-(c) denote different approaches to calculating p-values of the estimated treatment effects. All coefficients are statistically significant at the 1% level when OLS standard errors are used. However, clustering standard errors decreases the significance of all treatment effect estimates. The main columns of interest (1-3) remain significant at the 5% level. Finally, after the finite correction, the treatment effects measured in columns (1) and (3) remain statistically significant at the 5% level, and the coefficients of columns (2) and (5) become marginally significant.

Overall, the results indicate a positive treatment effect. In particular, being exposed to deep promotions during the first half of the experiment led consumers to be more likely to buy a promoted item during the second half, at which point they were faced with the same discount levels as their control counterparts. The intervention during the first half of the experiment induced a statistically significant increase in the probability of buying a promoted sku in 4.4 percentage points during the second half (column 1). Using statistic $E\left[\widehat{y}_i \middle| X_i, T_i = 1\right] \div E\left[\widehat{y}_i \middle| X_i, T_i = 0\right]$, this effect translates to a relative increase of 22.5%. Despite this, the quantity effect is relatively modest: columns (2) and (3) suggest that exposure to deep promotions led consumers to shift 1.6% of items in their baskets towards promoted goods, which translates into a 5.1% relative increase. The point estimates in

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

Table 6, columns (4)-(7), further suggest that the shift in basket compositions appears to source from an overall increase in the total number of purchases. However, we prefer not to emphasize these results, given the lack of statistical significance of the treatment effects in these columns.

Figure 6 decomposes the treatment effect across categories. It is constructed by setting the dependent variable to one (column 1 in Table 6) only when consumers make a purchase of a promoted good in each of the focal categories, and then performing the analysis at the category level. The point estimates range from 0.1% in the Tea Category up to 2.5% in the Milk category. All category treatment effects are positive, and near half are significantly different from zero, at the 5% level, with clustered standard errors corrected with the finite correction. None of the category results support the hypothesis that deeper promotions lead to consumer deal addiction.[17]

We test the robustness of our results in Appendix B. Given that retail settings are extremely busy with marketing stimuli, we perform the estimation on a sub-sample of consumers who are matched on observable attributes. This procedure is designed to reduce estimation noise, provided the consumers' unobservable attributes are correlated with the observable ones. The findings lend support to the hypothesis of heightened promotion sensitivity. Moreover, estimation precision increases, with the exception of the variables related to purchases of non-promoted goods (columns 6 and 7 in Table 6).

We investigate whether search can rationalize our findings as well as its implications for manufacturers in the next section. The role of search in retail environments has deserved recent attention in the marketing literature (see for example Seiler (2013) and Seiler and Pinna (2017)). Search behaviors are usually unobservable in retail settings. As such, the subsequent analysis should be taken as using the experimental findings to inform a model of search, rather than a proof that search behaviors underlie our results. However, it is useful to consider evidence consistent with the search mechanism, and to rule out alternative mechanisms.

By analyzing historical promotion depth patterns, we show that promotion depths are positively correlated across visits. This is a necessary condition for search behavior to ratio-

[17] 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.

nalize our findings: under consistent beliefs, a deep promotion offered by a brand today may lead consumers to search the same brand tomorrow, in hopes of finding a good deal once again. This result is discussed further in Section 4.3.

In terms of alternative mechanisms, we first investigate the extent to which the intervention may have generated state dependence in the measurement phase. State dependence, as operationalized by Guadagni and Little (1983) and further analyzed by Dube, Hitsch, and Rossi (2010), can be generated by several psychological mechanisms, including learning (Erdem, Imai, and Keane (2003)), consumers' thinking costs (Shugan, 1980) and consideration sets (Hauser and Wernerfelt (1990)).

In our setting, treated consumers were exposed to deep promotions during the intervention phase, and so bought more experimental goods than their control counterparts at the time. Hence, upon returning during the second half of the experiment, treated customers may exhibit more purchases of promoted goods because of the positive state dependence generated by the purchase behavior in the first half.

We show that state dependence cannot account for our results. In particular, all treatment estimates remain positive when the analysis focuses only on products bought for the first time during the measurement phase. In terms of statistical significance, the treatment effect on the likelihood of purchasing a promoted item becomes marginally significant, but additionally, all treatment effects remain statistically significance at the 95% level for consumers of the matched sample.[18]

Second, we perform a placebo analysis by analyzing consumers who did not buy any products in the 31 main categories during the first half of the experiment, and so are less likely to have been exposed to the treatment. All point estimates decrease toward zero, often by one order of magnitude. Only the treatment effect for the event of buying a promoted item remains marginally significant at the 90% level, providing a lower bound for the net treatment effect of 3.1 percentage points. Taken together, the results suggest that our main findings are robust to alternative explanations, in terms of magnitude and precision.

[18]The analyses and results are presented in Appendix B.

4 Searching for Deals

Having documented the experimental finding that consumers exhibit heightened promotion sensitivity, our goal is to recover the fundamentals for our observations under a behavioral model, and test its implications. This task implies considering a structural model of behavior that accounts for the potential forces in play. We use the experimental findings to investigate a model in which consumers search for deals. The model captures a few important features. First, it allows consumers to form beliefs about future promotion activities. In particular, the beliefs are informed by the historical promotion patterns at the retailer. Second, by allowing consumers to search for deals, it provides an explanation of consumer behavior that, together with the recovered beliefs, can rationalize our findings. Finally, it is used to assess the implications of a well-defined mechanism to firms competing through promotion activities.

4.1 Model

Consumer Search.

In order to derive implications about the phenomenon identified in the experiment, we consider a model in which consumers can search for deals before purchasing. We adopt the model of sequential search by Weitzman (1979), in which strategic consumers decide which products to evaluate, when to stop searching and which products to buy, if any. Evaluating alternatives is costly, and each consumer is required to incur search cost $c > 0$ to consider each one.

To illustrate the search model, consider the case in which a consumer decides whether to evaluate a single product, with unknown utility u_1, while having in hand an alternative that provides known utility u_0. If the consumer searches (subscripts i and t are omitted for clarity), she expects to earn

$$USearch_1\left(u_0\right) = -c + Pr\left(u_1 \geq u_0\right) E\left[u_1 \left| u_1 \geq u_0\right.\right] + Pr\left(u_1 < u_0\right) u_0 \tag{2}$$

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. Relatedly, note that a consumer with a high value of u_0 is less likely to search than one with a low value, ceteris paribus. Finally, there exists a level of u_0 that makes the consumer indifferent between searching alternative 1 or not. Define this threshold as the reservation value of alternative 1, denoted by u_1^*, which is implicitly defined by the solution to equation $USearch_1(u_1^*) = u_1^*$.

Rewriting equation (2) for some generic product j yields its implicit reservation value, given by

$$z_j^* = -c + \int_{z_j^*}^{\infty} udF_j(u) + F_j\left(z_j^*\right) z_j^* \tag{3}$$

where $F_j(\cdot)$ is a product-specific cumulative distribution function.

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

Consumer Beliefs.

Following the search literature, we assume consumers evaluate products in order to assess their idiosyncratic fit (e.g., products may be more useful at different occasions) as well as the prices they will have to pay to acquire them. Because in our setting promoted products are sold with special price labels made to stand out, we assume consumers can tell which products are sold in promotion easily. Moreover, in our dataset regular prices rarely vary, which informs the assumption that consumers know how much they will have to pay for products without special promotional labels a priori. Promotion depths do vary in our dataset however, and as a result consumer search can also be used to identify the discount levels offered by promoted products.

We use historical data over promotion depths, as well as consumer visit behavior, to inform beliefs. Noting that promotion depths vary over time, but are relatively highly correlated within each category during each promotional week, leads us to consider a category-level Markov process over category promotion depths, as depicted in Figure 1.[19]

[19] It is possible that consumers can track promotion depths at the product level, especially for products they buy regularly. While promotion depths are fairly highly correlated within products of the same category at each point in time, we simplify the estimation by considering category-level beliefs. We allow for beliefs at the individual product level in the counterfactual analysis. Finally, in the historical dataset, each category almost always has at least one product being promoted during in each promotional period. We do not use the

Figure 1: Discount Depth Transition Probabilities for a Category

$$
\begin{array}{cc}
 & \text{t:} \\
 & \begin{array}{cc} \text{Shallow} & \text{Deep} \end{array} \\
\text{t-1:} \begin{array}{c} \text{Shallow} \\ \text{Deep} \end{array} & \begin{bmatrix} 1-\omega^{S} & \omega^{S} \\ 1-\omega^{D} & \omega^{D} \end{bmatrix}
\end{array}
$$

Notation ω^{κ}, $\kappa \in \{S, D\}$, indicates the probability of a category being promoted with a deep discount in visit t, given that it was promoted with a κ-level discount (Shallow or Deep) in visit $t-1$. Under this specification, two consumers facing the same promotional activity may hold different beliefs over present discount depths, because of exposure to different promotional activities during their previous visits.[20] Empirically, we labeled discounts as shallow or deep based on the 15% cutoff, which falls between our experimental manipulations and also generates significant variation in the historical dataset.

Consumer Utility.

We assume consumer i derives utility from buying product j at purchase occasion t according to

$$u_{ijt} \;=\; \alpha_j + \beta SDep_{ijt} + \gamma^{D} d_{jt}^{Deep} + \gamma^{S} d_{jt}^{Shallow} + \varepsilon_{ijt} \tag{4}$$

where α_j captures the time invariant *net* utility of buying the product and paying its regular price. Regressor $SDep_{ijt}$ captures whether individual i purchased product j in the previous visit, and d_{jt}^{Deep} and $d_{jt}^{Shallow}$ are indicators for the discount depths of product j on occasion t. The product fixed effect α_j captures the *net utility* of buying the product and paying its regular price. This model feature stems from the facts that in our experiment regular prices are not manipulated, and moreover, we observe little variation in historical regular prices in the dataset, as mentioned before.[21] Finally, ε_{ijt} is an individual preference shock that

rare week-category combinations that do not feature any promotions in the estimation of the belief process.

[20] Index t is used to denote a consumer visit, except if noted otherwise.

[21] Because in our experiment regular prices were not manipulated, and moreover because we observe little variation in past regular prices, we assume that $p_{jt} = p_j$. In the model, $\alpha_j = \alpha_j^{'} - \beta_0 p_j$. Also, we simplify notation d_{jt}^{Deep} and $d_{jt}^{Shallow}$ to indicate the type of discount product j was sold with to consumer i during

captures the idiosyncratic fit of product j on occasion t. We assume consumers know the regular prices, which is consistent with our retail setting, in which regular prices rarely vary. Hence, when a consumer faces a product sold at a regular price, the taste shock is the only source of uncertainty. Consequently, expression (4) can be re-written as

$$u_{ijt}^{r} = v_{ijt} + \varepsilon_{ijt} \tag{5}$$

for products sold at regular prices, where $v_{ijt} \equiv \alpha_j + \beta SDep_{ijt}$ is the utility that consumer i knows she can derive from purchasing product j on occasion t.

Since promoted products are featured with salient labels that invite consumers to search in order to assess whether the deals are attractive enough, consumers can immediately identify the promotional status of products but are required to inspect pricing labels to assess promotional values. Hence, promoted products have an additional uncertainty layer induced by the discount depths. Consumer i's utility for a promoted product j is given by:

$$u_{ijt}^{p} = v_{ijt} + \varepsilon_{ijt} + \begin{cases} \gamma^{D} & with\, probability\;\; \omega_{ijt}^{\kappa} \\ \gamma^{S} & with\, probability\;\; 1 - \omega_{ijt}^{\kappa} \end{cases} \tag{6}$$

where $\kappa \in \{S, D\}$ denotes the promotional status of product j's category on occasion $t-1$.

We assume $\varepsilon_{ijt} \sim \mathcal{L}\left(0,1\right)$, where $\mathcal{L}\left(\cdot\right)$ is the Logistic distribution, and explain the benefit of this choice in the next section. Consequently, the product utilities of regular priced and promoted products, conditional on the consumer i's net utility from consumption and purchase history, follow distributions

$$u_{ijt}^{r}\Big|_{v_{ijt}} \sim F_{ijt}^{r} = \mathcal{L}\left(v_{ijt}, 1\right) \tag{7}$$

$$u_{ijt}^{p}\Big|_{v_{ijt}} \sim F_{ijt}^{p} = \omega_{ijt}^{H}\mathcal{L}\left(v_{ijt} + \gamma^{D}, 1\right) + \left(1 - \omega_{ijt}^{H}\right)\mathcal{L}\left(v_{ijt} + \gamma^{S}, 1\right) \tag{8}$$

Expression (7) states that, before search, the utilities from non-promoted products follow a logistic distribution. Expression (8) states that the utility of promoted products follows a mixture of Logistic distributions, induced by the unknown promotion depth consumers face before searching. Together with equation (3), distributions F_{ijt}^{r} and F_{ijt}^{p} determine the order in which alternatives are searched by consumer i on occasion t.

her t^{th} purchasing occasion.

Calculation of Reservation Values.

Estimation involves solving non-linear equation (3) for each iteration/consumer/alternative/occasion combination. Since the task is time consuming when tackled through numerical procedures, it is typically tackled by pre-computation of lookup tables of z_j^* as a function of c and v_{ijt}, under the assumption of normally distributed preference shocks (see Kim, Albuquerque, and Bronnenberg (2010) and Honka and Chintagunta (2017)). The procedure involves a startup cost but is efficient for estimation of parameters in typical search contexts. Below we motivate the choice of Logistic preference shocks, which yield closed-form expressions for the reservation values:

Proposition *(logistic uncertainty) The reservation value equation* (3) *admits a closed-form solution when $F_j(u)$ is Logistic with location parameter v_{ijt} and scale equal to 1, namely $z_j^* = v_{ijt} - \ln(\exp(c) - 1)$.*

The result above is useful to recover the reservation values of regular-priced products in our context. However, it is not helpful to solve for the reservation values of promoted products, which induce a mixed distribution. The calculation of the reservation values for the promoted products is performed by use of the following result:

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 $E(u|\, u > z)\, \forall z \in \mathbb{R}$.*

The contraction mapping procedure can be used to recover the reservation values for a broad class of distributions while controlling for numerical precision in a parsimonious way.[22]

We prove both of the results above in Appendix C. These results provide additional options for the estimation of search-based models. When researchers face no restrictions on the particular uncertainty specification, the logistic distribution assumption yields a convenient closed-form solution to the reservation values. In settings where preference uncertainty is

[22] It is important to note that all three methods are extremely efficient in calculating reservation values. Hence, the choice of method should depend on other factors, such as the specific search context and methodological convenience.

required to follow a different specification, however, the contraction mapping can be used to recover the reservation values efficiently for a general class of distributions. In our application we face both cases, and so the calculation of reservation values for regular-priced products takes advantage of the first result, and the calculation for promoted products takes advantage of the second one.

4.2 Estimation

Search Paths.

At the core of our investigation is consumer choice, which can be characterized according to

$$Pr\left(Choice_{ijt}|\theta, X_{ijt}, \omega_{ijt}\right) \tag{9}$$

where $Choice_{ijt}$ is the event that consumer i purchases product j on occasion t, conditional on some vector of preference parameters $\theta = \left\{\alpha_1..\alpha_J, \beta, \gamma^D, \gamma^S\right\}$, on some attributes $X_{ijt} = \left\{SDep_{ijt}, d_{jt}^{Deep}, d_{jt}^{Shallow}\right\}$, and on beliefs $\omega_{ijt} = \left\{\omega_{ijt}^S, \omega_{ijt}^D\right\}$. The object above is often investigated by use of a discrete choice model which, under typical assumptions, induces a logit likelihood that is especially amenable to empirical analysis. Since search paths are usually unobservable in retail settings such as ours, we integrate over the set $\mathcal{S}$ of all potential search paths that may have led to the consumer's choice, according to

$$Pr\left(Choice_{ijt}|\theta, X_{it}, \omega_{it}\right) = \sum_{S_{it}\in\mathcal{S}} Pr\left(Choice_{ijt}\left|S_{it}, \theta, X_{it}, \omega_{it}\right.\right) Pr\left(S_{it}\left|\theta, X_{it}, \omega_{it}\right.\right) \tag{10}$$

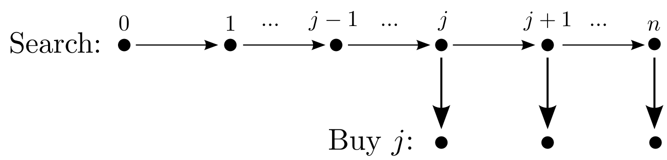
The expression above takes into account that, depending on the consumer's preferences θ over observables X_{it}, and on consumer beliefs ω_{it}, different search paths S_{it} may arise with different probabilities.

For illustration purposes, consider the case of a consumer deciding between two goods with uncertain utilities v_1 and v_2, and an outside option with known value v_0. Further, let product 1 have a higher reservation value than the second, $z_1 > z_2$, given some consumer preferences and beliefs. In this case, a consumer may opt for the outside option as a result of different search paths: she may decide not to search, to search only the first alternative, or to search both. She opts for the outside option in each of the following search sequences

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

The consumer selects the outside option with probability $Pr\,(Path\,1 \vee Path\,2 \vee Path\,3)$, and in general, calculating the probability of each option being selected involves adding over the search paths consistent with the respective choices. Figure 2 below shows the potential search paths consistent with a choice of some alternative j.

Figure 2: Search Paths Consistent with Choice of j



In Figure 2, searching an additional option corresponds to a lateral movement, and a downward one depicts the purchase of alternative j. In particular, 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 nonetheless worthwhile to continue to search, at least up to option j.

Identification and Estimation.

We start by discussing the identification of the preference parameters. We focus the discussion on the product-specific constants, α_j, since the identification of the preference parameters β, γ^D, and γ^S then follows, given variation in observables.

Conditional on a search cost and consumer beliefs, changes to an alternative-specific constant α_j only affect the reservation value of the same option j. Together with the assumption of mean zero valuation of the outside option, this fact implies a 1-1 correspondence with market shares.

As explained before, consumer beliefs are informed by historical promotion depth transitions, and are estimated in the first stage. As for the search cost, in the absence of search data it cannot be separately identified from the preference parameters. To see this, consider a market with two products, with market shares of 90% and 10%. Consumers may have an overwhelming preference for the first product because 1) it provides a much higher utility than the second one or 2) it is only slightly better than product two, but search costs are extremely high, such that most consumers prefer not to inspect the second product. Given the normalization imposed on the search cost, preference parameters should be interpreted as relative to consumers' propensity to inspect alternatives.

We faced a few estimation challenges: First, the integration across search paths implies taking a large number of simulations. Second, the likelihood function featured a number of saddle points, that made identifying the global maximum challenging. Third, imposing a normalization on the search cost can affect the ability of gradient-based methods to find the preference parameters that rationalize relatively similar market shares in the data. We solve these issues by employing a global patterned search across the parameter values. In addition, we employed the smooth estimator proposed by McFadden (1989) to calculate the Hessian of the log-likelihood, and the resulting standard errors. Our estimation procedure is performed in two stages. In the first stage we estimate the consumer beliefs that are consistent with the promotion patterns observed in the pre-experimental dataset. In the second stage we estimate the structural parameters conditional on the beliefs recovered in the first stage. The estimation details are provided in Appendix C.

We use both the intervention and the measurement phases to estimate the model at the consumer-visit level. The analysis focuses on the Milk category, as the implementation of the experiment closely followed the experimental design, as shown in Tables 3 and 5. Moreover, we focus on the behavior of customers of the matched sub-sample (see Section B.1) over both halves of the experiment, in order to improve estimation speed and efficiency. Within the Milk category, we model the actions of the largest three market share brands 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%) in the milk category and manage 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 rationalizes movements in intensive and extensive margins. However, we do not include occasions in which consumers bought no milk products, because promotions are likely to have a lower effect on overall consumption levels in this category.

Among the selected brands, brand 1 is the market leader and charges a price premium of 5% to 10% over the others, as shown in the first column of Table 7. Brand 2 is the retail chain's private label. Brand 3 holds the lowest market share and it is sold at an intermediate price point, as shown in the second column of Table 7.

4.3 Results

We recover the discount depth distribution from pre-experimental data to inform consumer beliefs of intertemporal promotion activities. We estimate the promotion depth Markov process at the category level using standard count methods. To compute transition probabilities, we assume an interpurchase time of 2 weeks, which accords well with the average time between visits that we observe in the data. The estimated transitions are later introduced into the search/purchase model at the individual level, depending on each consumer's visit patterns.[23]

We present the results of the discount-depth transition analysis in Table 8. 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, in all categories that exhibit a significant frequency of deep discounts, consumers are generally more likely to find deep discounts today after being exposed to deep discounts during their last visit (i.e. column 4 dominates column 2). This is consistent with the potential search explanation, that consumers may be more willing to evaluate discounted alternatives after being exposed to deep discounts in the past. In contrast, the search mechanism proposed here cannot predict heightened promotion sensitivity if deals were negatively intertemporal correlated: In this case treated consumers should be less willing to search for deals, during the measurement phase, than their control counterparts. In other words, the recovered beliefs do not pre-

[23]Our approach to recover consumer beliefs is consistent with most of the scanner panel literature. In contrast, imposing the Bayesian equilibrium concept 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.

clude the search mechanism as one of the explanations for our experimental findings. The recovered transition probabilities for the Milk category were incorporated into the model as consumer beliefs.[24]

Table 9 presents the estimates of the search model. As expected, 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 in the next purchase occasion. The point estimates of both discount levels are positive, with the high discount coefficient being larger than the shallow discount coefficient, as expected. Table 7 shows that the model recovers the market shares well, with discrepancies being less than two percent between predicted and actual market shares.

5 Counterfactual Analysis

5.1 Heightened Promotion Sensitivity

In order to test whether our search model can rationalize the experimental results, 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 facing the same discounts. We compute the difference in purchase probabilities between consumers who expect deep discounts relative to consumers who expect shallow discounts, despite facing the same shallow discount realization, i.e.,

$$Pr\left(Buy_j|\,\omega^D, d_j^{Shallow}=1\right) - Pr\left(Buy_j|\,\omega^S, d_j^{Shallow}=1\right) \tag{11}$$

We incorporate consumer beliefs in the following way. The first term of equation (11) corresponds to the probability of buying brand j, given that the consumer faced deep discounts in the past, and hence expects a shallow discount with 55.6% probability (according to Table 8, column 3). Similarly, the second term of equation (11) corresponds to the probability of

[24]Because our intervention is relatively short, it is unlikely that if affected consumer beliefs. Hence, the recovered historical promotion transitions are used as the beliefs held by consumers for estimating the model, as well as in the counterfactual analysis.

buying brand j given that the consumer faced a shallow discount for milk in the past, and hence expects a shallow promotion with 83.9% probability (Table 8, column 1).

The results are summarized in Table 10. They imply that the differential beliefs about promotion depths, as informed by the historical dataset, lead to a relative sales increase of 14.4% for promoted products, which can be compared with the result of 21.2% found in the analysis that takes state dependence into account (Appendix B). The underlying intuition is that, on the margin, consumers prefer to search products they believe are more likely to feature deep promotions. These products have the advantage of being considered first, and so market share gains result.

5.2 Implications for Competitors

In this section we investigate the implications of consumer search for deals, as proposed by our model, for competing manufacturers. 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 only need to offer shallow discounts for a while to generate positive results on sales. This interpretation is consistent with the previous counterfactual analysis, in which firms are believed to offer discounts unilaterally. We now consider the interaction of manufacturers who compete for consumers through promotion activities.

Supply-side Model.

We consider a two-period model in which firms compete through promotions strategically. Their action space is to sell at the regular price (i.e. no promotion), offer a shallow discount or a deep discount. Formally, in period $t \in \{1, 2\}$, firm $j \in 1...n$ 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 relevant for competition. In particular, vector Ω_t contains a sub-vector $\widehat{\omega}_t$ $(n \times 1)$ that summarizes consumers' expectations about the discount levels, and an additional vector with consumer state dependence information, related to the previous period's purchases. The first period's states, Ω_1, are initialized by the researchers according to the brands' market shares in the data.

The demand side is informed by the representative consumer's parameters estimated in the previous section. Each firm's objective is to maximize its expected two-period profit.[25] Equilibrium profits are given by

$$\Pi_j^* = \pi_{j1}^* + \pi_{j2}^* \tag{12}$$

where, in particular, π_{jt}^* is defined as

$$\pi_{jt}^* = \max_{Disc_{jt} \in \{None, Shallow, Deep\}} E\left\{ \left(p_j \left(1 - Disc_{jt}\right) - mc_j\right) \cdot D_j \left(Disc_{jt}, Disc_{-jt}^*, \Omega_t \right) \right\} \tag{13}$$

where p_j and mc_j are the firm's price and marginal cost, respectively. The regular prices in the data inform p_j, whereas we simulate the model at several degrees of marginal costs in order to incorporate the fact that price promotions may sometimes be directly linked to manufacturer-related conditions, including changes in opportunity costs of not selling.[26]

We focus on subgame perfect equilibria, such that firms understand that their actions influence future market conditions as well as the future actions of their competitors. The model is solved by backward induction. In addition, we use the fact that the firms' actions in the first period are a sufficient indicator to describe state vector Ω_2.[27]

Once the second period payoffs are calculated for each of the first-period action profiles, a deviation analysis is performed in order to isolate potential equilibria of the second subgame. Tentative equilibrium path profits are then plugged into the first period optimization problem, where all action profiles are re-visited. We are able to identify all pure subgame perfect equilibria by inspecting all possible first-stage action profiles. When a pure-strategy equilibrium does not exist, an exhaustive search is employed to uncover mixed strategy outcomes.[28]

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

[26] In our context, most promotions are initiated by manufacturers. We assume 100% passthrough in the analysis, although the effective rate is slightly lower. While we condition on different opportunity costs, in reality they are likely to change over time, leading to different equilibrium promotion profiles in different weeks.

[27] Theoretically, firms could observe individual consumers' buying behaviors, and hence have a precise idea of the realized state dependence in the market at the individual level. In reality, firms may not always take all consumer behavior into account, and thus have to integrate over the distribution of state dependence in order to optimize their promotional offerings. In our case, our homogeneous demand structure allows us to summarize state dependence effects through past market shares.

[28] Given the number of players and the size of the action space, finding mixed strategy equilibria can be tedious. Whenever a pure strategy equilibrium was not found, our estimation code prepared a file to be read by the Gambit software (www.gambit-project.org), which was then used to run all algorithms available to generate an exhaustive list of mixed strategy outcomes. These algorithms always found a unique mixed-strategy outcome.

Simulation.

We compare the market competition with and without heightened promotion sensitivity. In the first scenario consumers search according to the recovered demand-side estimates and recovered beliefs for the Milk category. In the second scenario, consumers expect shallow discounts throughout, and never adapt their beliefs about promotion activities. Specifically, in the first scenario consumers share the same initial belief of facing shallow discounts with 83.9% chance, but decrease it to 55.6% in the second period for firms who offer deep discounts in the first period. In contrast, in the second scenario consumers expect shallow discounts occur with probability 83.9% in both periods, independently of firms' promotional activities. Hence, the baseline beliefs are constant across the scenarios, but firms are able to affect consumer beliefs in the first one.

Firms can sell their products at regular prices, equal to the ones in the dataset, at shallow 4.4% discounts or at deep 23.9% discounts, in accordance to the discounts effectively implemented in the experiment, as reported in Table 3.

We characterize market outcomes at a wide range of marginal costs in order to contemplate the following cases: 1) in some circumstances retailers want to promote specific products, making the manufacturer's relevant counterfactual margin higher than just the strict contribution margin; 2) manufacturers face inventory and cost variations over time as well as expiration dates, all of which affect their opportunity cost of selling. Such forces can lead accounting margins to underestimate economic ones. Broadly speaking, in the counterfactual analysis, a higher margin can be interpreted as a higher incentive to offer promotions.

Table 11 presents the subgame perfect equilibrium action profiles at different margin levels, when consumers exhibit heightened promotion sensitivity, or alternatively, use promotion transitions to form their beliefs. In the first period, firms prefer to engage in shallow promotions, up to the 40% margin level. From the 50% level onward, firm 3 transitions to offering deep promotions. This is in contrast with the second period, in which all firms prefer shallow promotions. The reason for this difference is that, in the first period, offering deep promotions can be beneficial because firms can then take advantage of state dependence in the second period. The policy change by firm 3 is intuitive in the sense that it is the firm with the lowest market share, and as a result, the one with the most to gain in inviting consumers to include it in their search paths.

When margins are high (80%), or alternatively, when opportunity costs of not selling are very high, the effective cost of offering deep promotions is lower, and so all firms prefer to offer deep promotions, albeit sometimes only probabilistically: at the 80% level, firms 1 and 2 engage in mixed strategies between shallow and deep promotions.

Table 12 presents the same figures for the case in which consumers do not exhibit heightened promotion sensitivity. As before, all firms offer shallow discounts in the second period. In the first period, however, they prefer to offer discounts in fewer conditions. Firm 3 only offers deep discounts starting at the 60% margin level, and the remaining firms only begin offering deep discounts, probabilistically, at the 90% margin level. The reason for the increased reluctance to introduce deep discounts, when compared to the scenario of heightened promotion sensitivity, is that in this case consumers do not expect a higher likelihood of a firm promoting in the second period as a function of its promotional activity in the first period. As a result, current promotions no longer provide an incentive for consumers to search in the second period, and so the returns of offering deep discounts in the first period become lower.

Table 13 shows the profit ratios between both scenarios, where a number above 1 means that the firm makes higher profits in the heightened promotion sensitivity scenario. In the 50%-70% 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 it is the one that gains the most from inducing search by consumers through discounts. Firm 3's discount has a -4% profitability impact on firm 1, at the 50% margin, which then attenuates at higher margin levels. At 80% margins, firms 1 and 2 offer deep discounts with positive probability. The result is that all firms become worse off because of heightened promotion sensitivity. Cabral and Villas-Boas (2005) have coined this type of result as a Bertrand supertrap, in which an apparent advantage for a monopolist ends up effectively decreasing all firms' profits due to fiercer competition.

The results indicate that, in competition, promotional pressure sources from firms with smaller market shares first. Moreover, firm 3's gains from dynamic beliefs are non-monotonic in contribution margins: 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 promotion 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, with small firms losing the most.

6 Conclusion

We implemented a large scale field experiment that involved changing the prices of 170 products across 17 categories in 10 supermarkets, in order to investigate the dynamic effects of price promotions in a retail setting. We found that deep promotions heighten customers' future promotion sensitivities. In particular, customers were 22.5% more likely to buy promoted goods after being exposed to 30% discounts rather than 10%, ceteris paribus. Along the same lines, the proportion of promoted goods in consumers' baskets increased in 5.1% for treated consumers relative to their control counterparts. The point estimates of the treatment effects are positive across all categories.

We use a matching procedure to verify the robustness of our results. Consistent with econometric theory, the statistical significance of the treatment effect increases, while keeping most treatment effect estimates statistically similar. The procedure also provides statistical power to separate the experimental effects from those generated by state dependence. Finally, a placebo test shows no significant effects on consumers who did not make a purchase from a major category during the first half of the experiment.

Finally, we investigate a model in which consumers search for deals, and show that, conditional on the historical promotion patterns observed at the retailer, it can rationalize our experimental findings. The underlying intuition is that when consumers believe that promotion depths are positively correlated over time, offering a promotion today can be viewed as an invitation to consider the same product tomorrow. The counterfactual analysis documents that as the incentive to offer promotions increases, smaller brands with lower market shares have a higher incentive to offer promotions in order to invite consumers to consider them.

Our results provide a rationale for why some managers complain about competing on promotions while at the same time we observe deep promotions being routinely offered in equilibrium. Since consumers exposed to deep promotions display heightened promotion sensitivity, and furthermore, since firms can induce greater subsequent purchases by offering deep promotions, competition on the intensity of promotional activities can hurt firms' profitability: firms may find themselves in a prisoner's dilemma in which they compete fiercely on promotional activities.

In terms of future research, it may be interesting to consider how heightened promotion sensitivity affects firms in the long run, as well as the effects on retailer profitability. Although our model does not focus on the retailer's role, by assuming full passthrough, an implication is that the retailer can use passthrough as a coordination device that allows firms to move to an outcome characterized by fewer intensive promotional activities and greater overall profitability. Since larger margins provide an incentive for firms to offer deeper promotions, it is likely that the retailer can induce firms to reduce their reliance on deep promotions through lowering the extent to which trade promotions are passed through to retail prices. Counterintuitively, having the retailer pay lower wholesale prices can lead to higher manufacturer profits due to softer competition.

Tables

Table 1: Descriptive Statistics by Week-Store

	Mean	Std. Dev.	Minimum	Maximum
Number of Visits	10,267.6	3,742.7	3,239	17,323
Total Sales (USD)	$166,517.7	$89,057.1	$54,874.7	$348,648.4
No. of Items Sold	70,636.7	33,588.7	20,218	130,831
- Experimental	1,707.3	944.7	375	4,548
- Non-experimental	68,929.5	32,735.2	19,783	127,259
Average Basket Size (items)	6.7	1.2	4.8	9.9
Number of Weeks	5			
Number of Stores	10			

Notes: The figures above are calculated during the second half of the experiment, for the selected sub-sample of 234,063 consumers, across experimental categories (excludes Candy, Cheese, Cooking Oil, and Meats categories).

Table 2: Descriptive Statistics by Experimental Condition

	Treated Stores (1)	Control Stores (2)
Demographics		
- Age	47.26 (14.26)	45.68 (14.27)
- Fraction Female	.657 (.475)	.648 (.478)
Pre-experimental Expenditure (USD per visit)		
- Total Expenditure	$66.70 (60.86)	$55.43 (73.55)
- Expenditure on Experimental Categories	$20.44 (21.43)	$17.20 (22.16)
- Total Number of visits in 46 weeks	30.28 (12.92)	26.32 (12.97)
Number of customers	115,129	106,194

Notes: Standard deviations in parentheses. Pre-experimental data is available only for 221,323 out of the 234,063 individuals considered in the pooled regressions.

Table 3: 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.

Table 4: 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.

Table 5: Price and Promotion Differences across Experimental Phases

	% Change in regular prices		% Change vs. promotions 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 3 and Table 4.

Table 6: 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	0.044	0.016	0.017	0.255	0.295	0.707	0.672
(a): OLS std. errors	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
(b): Clustered std. errors	(0.014)*	(0.021)*	(0.011)*	(0.088)†	(0.061)†	(0.074)†	(0.081)†
(c): (b) + bootstrap-t	(0.048)*	(0.061)†	(0.044)*	(0.135)	(0.093)†	(0.102)	(0.113)
Constant	0.308**	0.218**	0.206**	1.383**	1.475**	4.871**	5.569**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.066	0.002	0.0025	0.056	0.054	0.096	0.105
N. Observations:	234,063	122,762	122,762	234,063	234,063	234,063	234,063

Notes: † $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$. 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 7: 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: Actual market shares computed from the data. Predicted market shares obtained through the structural model.

Table 8: 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 9: 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	26,964	
N. Alternatives	3 per choice occasion+outside option	

Notes: * p-value $\leq$ 0.01. Standard errors in parentheses.

Table 10: Counterfactual Analysis: Heightened Promotion Sensitivity

	$Pr\left(Buy_j \vert \omega_{ic}^S, d_j^{Shallow}=1\right)$	$Pr\left(Buy_j \vert \omega_{ic}^D, d_j^{Shallow}=1\right)$	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 11: Competition Counterfactuals: Heightened Promotion Sensitivity

- 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 Expected payoffs, per *potential* customer (assumes at most 1 purchase per customer), 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 12: Competition Counterfactuals: No Heightened Promotion Sensitivity

- 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 Expected Payoffs, per *potential* customer (assumes at most 1 purchase per 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 13: Firms' Profitability Implications of Heightened Promotion Sensitivity

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 of the firms' profits under the assumption that consumers exhibit heightened promotion sensitivity vs. not.

Figures

Figure 3: Timing of Discounts

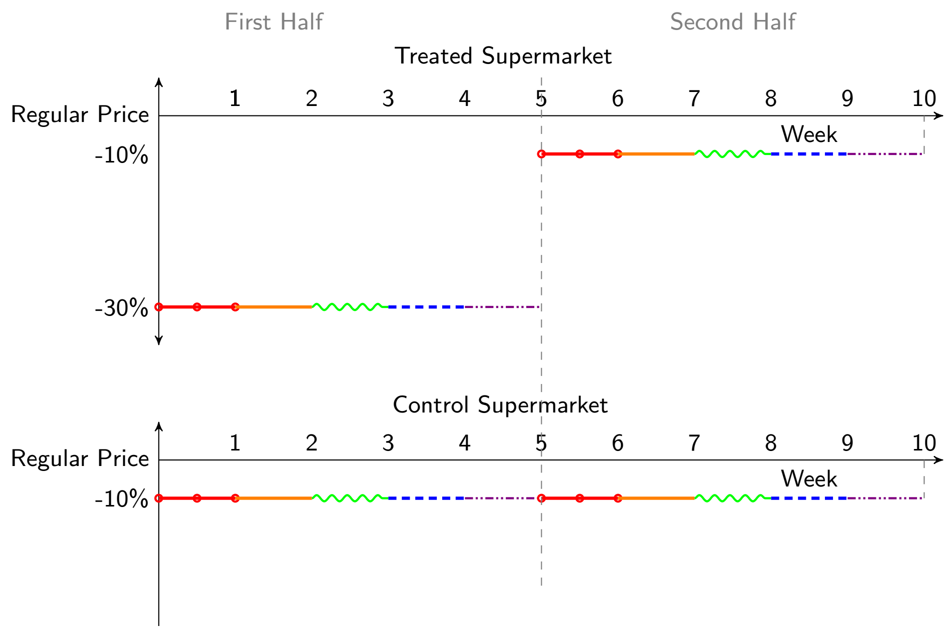


Figure 4: Pre-experimental Shopping Behavior across Experimental Conditions

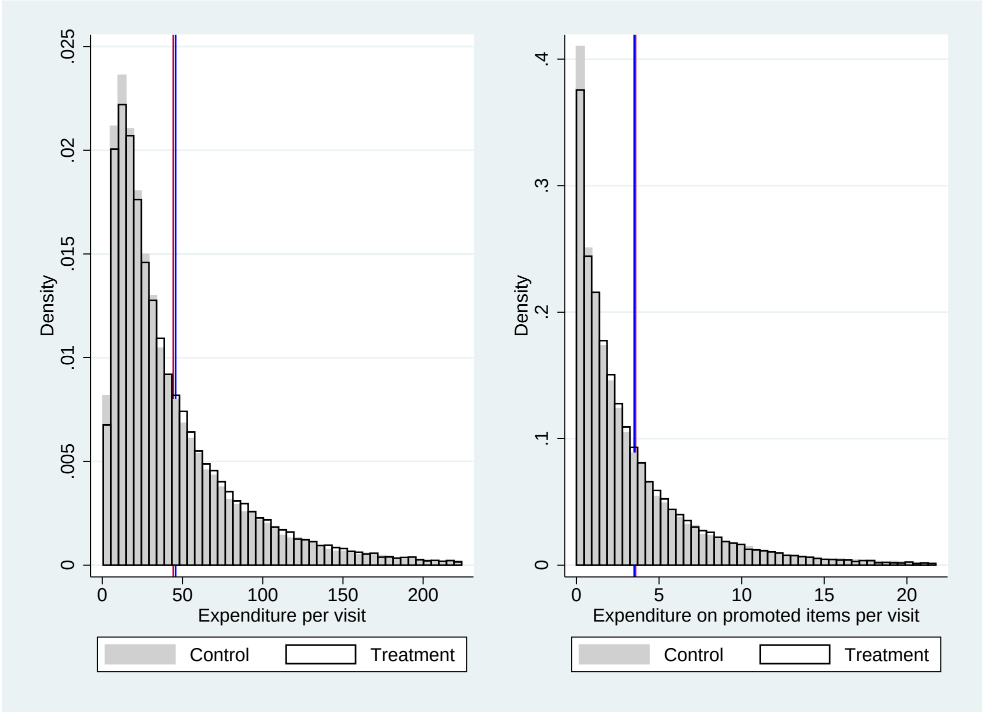
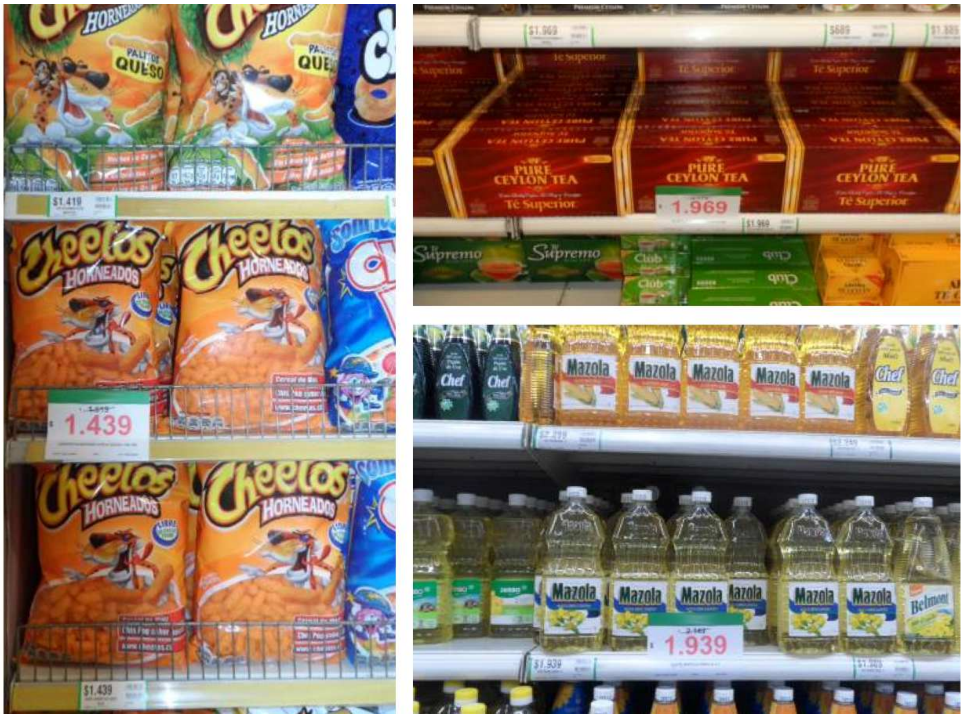
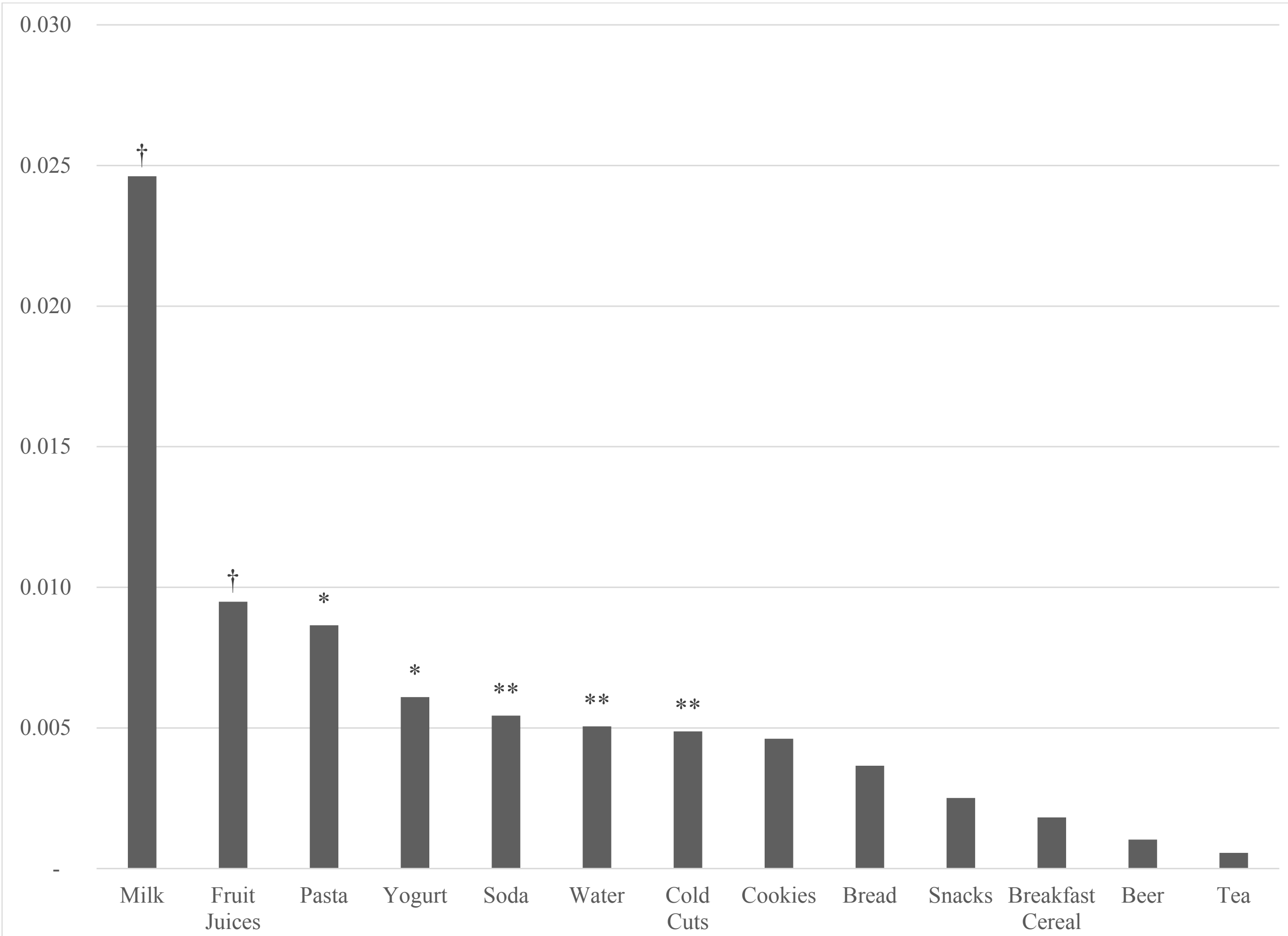


Figure 5: Experimental Promotions in the Retail Space



Clockwise: Experimental Promotions in the Snacks, Tea and Cooking Oil categories.

Figure 6: Category-Level Treatment Effects



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 10,000 draws. The analysis is based on matched customers. N=234,063 in all regressions.

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 could 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 Robustness/Alternative Mechanisms

B.1 Improving Efficiency Through Matching

To improve the efficiency of our treatment estimates, we complement our analysis 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 experimental period (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 level.[29] Notice 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.[30]

Matching Framework.

To construct our sample of statistical twins, we introduce a recent matching technique developed by Zubizarreta (2012). This matching technique takes advantage of new developments in optimization to match individuals in multiple dimensions, which until recently was an unfeasible task due to the large 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 treated unit t is assigned to control c, and 0 otherwise; and denote the entire assignment

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

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

matrix by $\mathbf{a}$.[31] The optimal assignment problem is given by:

$$\min_{\mathbf{a}} \sum_{t\in\mathcal{T}} \sum_{c\in\mathcal{C}} \delta_{t,c} a_{t,c} \tag{14}$$

$$\text{subject to} \sum_{c\in\mathcal{C}} a_{t,c} = 1 \quad , \quad t \in \mathcal{T} \tag{15}$$

$$\sum_{t\in\mathcal{T}} a_{t,c} \leq 1 \quad , \quad c \in \mathcal{C} \tag{16}$$

$$a_{t,c} \in \{0,1\} \quad , \quad t \in \mathcal{T}, c \in \mathcal{C} \tag{17}$$

$$\left| \sum_{t\in\mathcal{T}} \sum_{c\in\mathcal{C}} \frac{x_{c,j} a_{t,c}}{T} - \overline{x}_{\mathcal{T},j} \right| \leq \varepsilon_j \quad , \quad j \in \{1,..,P\} \tag{18}$$

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 (14). The first three constraints describe the integer nature of the assignment problem: Each treatment unit is paired with one control unit (Equation (15)) and not all control units should be used (Equation (16)).[32] The set of constraints given by expression (18) 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 the mixed-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. We then applied the criteria that customers are required to buy an item in at least one of

[31] 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.

[32] 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.

the 31 main categories during the first half of the experiment and not visit multiple stores during the experimental period. The historical dataset used for this task covers a 46 week period, one year prior to the experiment.

Table 14 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 $13,482$ one-to-one customer pairs (one control and one treated customer), distributed across 10 stores of our retail chain. Table 15 presents the resulting covariates of the final matched sample.[33] 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 analyses were 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.

Matched Sample Analysis.

We re-estimate equation (1) 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.[34]

Table 16 presents the results using the matched subsample to confirm that consumers display heightened promotion sensitivity. The first column presents the effect of exposure

[33]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.

[34]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.

to deep promotions on the purchase incidence of promoted products. Treatment induces an increase in probability of buying a promoted sku in 5.5 percentage points, and is statistically significant at the 95% confidence level. In relative terms, consumers exposed to deep promotions are 21.2% more likely to buy promoted items than their control counterparts, despite facing similar promotional levels. As before, the second and third columns also suggest a shift in consumers' basket compositions towards promoted products. Importantly, the significance of all treatment effects of interest (columns 1-3) increases.

Table 14: Universe of Potential Pairs

Sample:	Before Matching		After Matching		
	(1)	(2)	(3)	(4)	(5)
Store Pair	Treatment	Control	Treatment	Control	Sample
1	32,711	8,535	1,647	1,647	3,294
2	20,477	30,746	5,647	5,647	11,294
3	16,035	39,773	4,116	4,116	8,232
4	20,230	17,497	1,322	1,322	2,644
5	16,741	18,578	750	750	1,500
Total	106,194	115,129	13,482	13,482	26,964

Notes: Individuals matched by a Mixed Integer Programming procedure (Zubizarreta (2012)) using pre-experimental data. Pre-experimental data is available only for 221,323 out of the 234,063 individuals considered in the pool regressions.

Table 15: Pre-treatment Covariates of Control and Treated Matched Individuals

Pre-Treatment Covariate	Control	Treatment	Difference	p-value
Average Weekly Total Expenditure (USD)	\$79.71	\$80.08	-\$0.38	(0.14)
Average Weekly Expenditure on Experimental Categories	\$26.09	\$26.31	-\$0.22	(0.30)
Age	47.26	47.33	-0.07	(0.07)
Fraction Female	0.66	0.64	0.03*	(0.00)
Total Number of visits in 46 weeks	27.27	28.11	-0.83*	(0.00)

Sample Size: 26,964

Comparing the point estimates in Table 6 to those in Table 16 we verify that, except for the coefficient of the proportion of promoted products, all coefficients from the regression using the matched sample fall within the confidence intervals of the coefficients using the full sample. This is suggestive evidence that the matching procedure did not alter the distribution of treatment effects in a statistically significant way.[35]

[35]Importantly, the treatment effects estimated on the matched sample remain valid even if the matched consumers behave differently from the ones in the full sample. 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 finite sample correction implies an impractical amount of computation time.

Table 16: 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	0.055*	0.037*	0.034*	$0.359^{\dagger}$	$0.346^{\dagger}$	0.713	0.653
	(0.023)	(0.039)	(0.036)	(0.082)	(0.071)	(0.123)	(0.119)
Constant	0.299**	0.211**	0.196**	1.343**	1.463**	4.919**	5.6**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.068	0.006	0.006	0.054	0.051	0.092	0.107
N. Observations:	26,964	10,022	10,022	26,964	26,964	26,964	26,964

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.

B.2 State Dependence

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%). It is possible that, during the second half of the experiment, both customers visited the store on the week that the same product was on promotion once again, 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.

We repeat the main analysis, but now only consider, for each consumer, purchases of goods that were not bought during the first half of the experiment. While this procedure is expected to mechanically decrease the treatment effect due to ignoring relevant data, it has the merit of parsing out the effect of state dependence. Table 17 summarizes the results: the treatment effect in column (1) becomes marginally significant, and all treatment effect estimates decrease. However, statistic $E\left[\widehat{y}_i \middle| X_i, T_i = 1\right] \div E\left[\widehat{y}_i \middle| X_i, T_i = 0\right]$ produces an estimate of a 21.2% relative increase of purchases of promoted products, which is similar to the one of 22.5% found before, when state dependence was not controlled for.

In order to investigate this issue further, we consider the same analysis on the matched sample, discussed in the previous sample, and present the results in Table 18. All results of interest (columns 1-3) remain statistically significant in this case, with the treatment estimates falling only slightly. Taken together, the results imply that state dependence may play a role in our measurement, but is unlikely to be responsible for the finding of heightened promotion sensitivity.

Table 17: 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	$0.037^{\dagger}$	$0.016^{\dagger}$	0.018*	0.148	$0.19^{\dagger}$	0.348	0.325
	(0.066)	(0.062)	(0.023)	(0.134)	(0.067)	(0.117)	(0.162)
Constant	0.235**	0.223**	0.215**	0.838**	0.961**	2.81**	3.325**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.04	0.002	0.002	0.03	0.029	0.06	0.06
N. Observations:	234,063	112,043	112,043	234,063	234,063	234,063	234,063

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 18: Effect of Treatment on Customer Behavior for New Purchases - 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	0.048*	0.04*	0.033*	$0.227^{\dagger}$	$0.218^{\dagger}$	0.361	0.32
	(0.026)	(0.025)	(0.023)	(0.057)	(0.069)	(0.125)	(0.179)
Constant	0.227**	0.214**	0.197**	0.824**	0.992**	2.839**	3.404**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.04	0.006	0.005	0.03	0.024	0.056	0.06
N. Observations:	26,964	8,330	8,330	26,964	26,964	26,964	26,964

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.

B.3 Stockpiling Behavior

Our results are likely to be made conservative by consumer stockpiling behavior. The reason is that treated consumers are likely to hold higher inventories than control ones, due to the exposure to deep promotions during the intervention phase. As a result, stockpiling is expected to dampen treated consumer purchases during the second half of the experiment, including those of promoted products.

B.4 Placebo Test

In this section we introduce a placebo test designed to assess whether our experimental intervention is likely to be effectively 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 19. 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. While the Placebo test is diagnostic, it is likely to be at least partially contaminated to different extents by two forces with opposing implications. First, the choice of not purchasing from a major category during the first half of the experiment is unlikely to be exogenous. Because of this, the lack of a statistically significant treatment effect can source from selection, i.e., the test just so happens to select few consumers who respond to price promotions in general. A countervailing argument is that it is impossible to rule out that these customers were not exposed to our intervention as well, in which case the lack of statistically significant treatment effects is more reassuring.

Table 19: 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 non-promoted sku's bought	(7) Expenditure on non-promoted sku's
Treatment	0.013†	0.006	0.003	0.055	0.033	0.128	0.1
	(0.086)	(0.618)	(0.805)	(0.28)	(0.506)	(0.478)	(0.484)
Constant	0.105**	0.208**	0.193**	0.321**	0.378**	1.289**	1.713**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Controls & Store Group Fixed Effects	✓	✓	✓	✓	✓	✓	✓
R-Squared (Within)	0.005	0.007	0.007	0.008	0.005	0.014	0.014
N. Observations:	9,059	2,594	2,594	9,059	9,059	9,059	9,059

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 50,000 draws. Columns 4 and 5 only include treated-control pairs who have bought items during the second half of the experiment.

C Search Model

C.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 $\mu, c \in \mathbb{R}$.

C.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 $\Gamma(z) = -c + \int_z^\infty vf(v)\,dv + zF(z)$. Under standard continuity assumptions, $\Gamma(z)$ is a contraction mapping if $\Gamma'(z) \in [0,1)$. In our case,

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

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 $\int_{z_n}^\infty xf(x)\,dx$ is finite $\forall z_n \in \mathbb{R}$, which is also implied by the original Weitzman (1979) model. So, the theorem applies to most distributions used in empirical work.

Using the proposition above, it is easy to show that in the case of mixture of logistics given by equation (8), 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 $H \in \{S, D\}$.

C.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 2, 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. 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

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

We now characterize each of the paths, where movements referred to as 'down' and 'right' are related to the ones in the Figure 2:

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

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 $\eta \sim N(0,1)$ 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 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.[36] Moreover, the log-likelihood function exhibits saddle points that make finding the global maximum challenging.

We implement a patterned grid search across a wide range of parameter values, and ensure that the bounds set 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

[36] See Train (2009) (Sec. 5.6.2) for a careful exposition of this issue.

component.[37] 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

$$p^r\left(Choose_{n-1}\right) = K\left(\min\left\{z_{n-1}, v_{n-1}^r\right\} - \max\left\{v_0^r..v_{n-2}^r\right\}\right).\left[K\left(v_{n-1}^r - z_n\right) + K\left(z_n - v_{n-1}^r\right).K\left(v_{n-1}^r - v_n^r\right)\right]$$

where

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

is the logistic kernel with smoothing parameter $\sigma = 0.001$. 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.

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

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

[37] See Honka and Chintagunta (2017) for an application within the search framework.

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.

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.

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.

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.

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.

Kopalle, P. K., C. F. Mela, and L. Marsh (1999): “The dynamic effect of discounting on sales: Empirical analysis and normative pricing implications,” *Marketing Science*, 18(3), 317–332.

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.

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.

Seiler, S., and F. Pinna (2017): "Estimating Search Benefits from Path-Tracking Data: Measurement and Determinants," *Marketing Science*, Forthcoming.

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*.

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.

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.