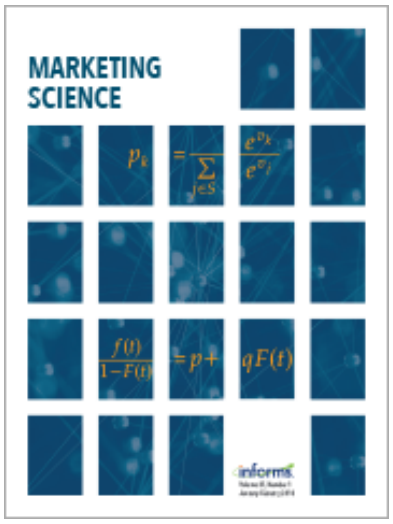


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Long-Run Effects of Promotion Depth on New Versus Established Customers: Three Field Studies

Eric T. Anderson, Duncan I. Simester,

To cite this article:
Eric T. Anderson, Duncan I. Simester, (2004) Long-Run Effects of Promotion Depth on New Versus Established Customers: Three Field Studies. Marketing Science 23(1):4-20. https://doi.org/10.1287/mksc.1030.0040

MARKETING SCIENCE
Vol. 23, No. 1, Winter 2004, pp. 4–20
ISSN 0732-2399 | EISSN 1526-548X | 04 | 2301 | 0004

informs

DOI 10.1287/mksc.1030.0040

Long-Run Effects of Promotion Depth on New Versus Established Customers: Three Field Studies

Eric T. Anderson
Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, Illinois 60208, eric-anderson@kellogg.northwestern.edu

Duncan I. Simester
Sloan School of Management, Massachusetts Institute of Technology, 38 Memorial Drive, Cambridge, Massachusetts 02142, simester@mit.edu

We use the results of three large-scale field experiments to investigate how the depth of a current price promotion affects future purchasing of first-time and established customers. While most previous studies have focused on packaged goods sold in grocery stores, we consider durable goods sold through a direct mail catalog. The findings reveal different effects for first-time and established customers. Deeper price discounts in the current period *increased* future purchases by first-time customers (a positive long-run effect) but *reduced* future purchases by established customers (a negative long-run effect). Overall, the results show evidence of several long-run effects: forward buying, selection, customer learning, and increased deal sensitivity. Short-run metrics that ignore these effects overstate the overall change in demand for established customers. The implication is that if prices are set based on short-run elasticity, then they will be too low. Among first-time customers, the short-run metrics underestimate the total increase in demand. If prices are set based on short-run elasticity, then they will be too high.

Key words: price promotions; pricing; long-term effects; forward buying; purchase acceleration; deal sensitivity; catalogs
History: This paper was received February 19, 2002, and was with the authors 10 months for 2 revisions; processed by Scott Neslin.

1. Introduction

We present the results of three large-scale field experiments that show that the depth of a price promotion can affect repeat-purchase probabilities even up to two years later. The three studies were conducted with the cooperation of a mail-order catalog firm that sells durable goods. A key feature of the studies is that we are able to separately identify customers who have never purchased from this catalog before (prospective customers) from customers who have purchased before (established customers). With respect to repeat-purchase rates, we find that increasing the depth of a price promotion has a positive long-run effect among prospective customers and a negative long-run effect among established customers.

Of the three studies we present, two involved prospective customers, while the third study involved established customers. In each study, either a Promotion or a Control version of a "Test Catalog" was sent to randomly assigned samples of actual customers. The two versions were identical with the exception that the depth of the price promotion was greater for some items in the Promotion version. Following the Test Catalog, there was no difference in the subsequent catalogs sent to customers in the two conditions. To measure the long-run impact of promotion depth, we tracked repeat purchases of customers who responded to the Test Catalog for a period of approximately two years.

There are at least two reasons to anticipate fewer repeat purchases (on average) among customers who purchased from the Promotion version of the Test Catalog versus those who purchased from the Control version. First, lower prices in the Promotion version may prompt customers to *forward buy* (purchase acceleration) to meet future needs (Krishna 1992, 1994). Second, lower prices in the Promotion version may prompt purchases from the Test Catalog by customers with lower product valuations (Neslin and Shoemaker 1989), which we will refer to as *selection*. Both selection and forward buying could contribute to fewer repeat purchases in the Promotion condition, even though all customers were mailed the same

future catalogs. The findings confirm this prediction, but only with established customers. Among customers who purchased for the first time from the Test Catalog, future demand was higher in the Promotion condition. This led to two benefits for the firm: Deeper promotions attracted more first-time customers and led to higher repeat-purchase rates for these first-time customers.

Overall, our findings confirm that the depth of a price discount can have important effects on future demand. They also highlight the importance of distinguishing between first-time and established customers. Given that the impact on repeat purchase rates is positive for one segment and negative for the other, failure to distinguish between the segments would lead to misleading estimates of the long-run effects. This has important implications for firms' pricing strategies. If firms focus solely on short-run elasticities, or they fail to distinguish between first-time and established customers, then prices may be set incorrectly. Among established customers prices may be too low, while among first-time customers they may be too high.

While there is a vast literature on price promotions, evidence of long-run effects has been mixed. In a review article, Blattberg et al. (1995) conclude that the long-run effect of price promotions is "probably the most debated issue in the promotional literature and one for which the jury is 'still out'," (p. G127). Recent studies using either household or store-level data from the packaged-goods industry have yielded mixed results. Studies using household data report negative effects (Mela et al. 1997, 1998; Jedidi et al. 1999), no effect (Pauwels et al. 2002), and positive effects (Ailawadi and Neslin 1998). Two recent studies using store-level data report no permanent long-run effect of price promotions (Dekimpe et al. 1999, Nijs et al. 2001).[1]

Some of the confusion over whether there are long-run effects rests in the definition of "long run." Our studies take place over a period of approximately two years, whereas several of the previous studies take place over a much longer time horizon (Mela et al. 1997, Nijs et al. 2001). In these studies, a time period of two years might be defined as the "medium term" rather than "long term." We note that the average interpurchase time in our studies is 48 weeks and leave the interpretation of whether the two-year time horizon in our studies is "medium term" or "long term" to the reader.

[1] These papers differ somewhat in their definition of a long-run effect. The vector autoregression (VAR) time-series literature refers to a permanent effect as not mean reverting, implying a change that lasts forever. The distributed lag response models view long run as quarters or years, but allow for mean reversion.

In our studies, we find evidence of both positive and negative long-run effects, but our approach differs considerably from previous work. First, these previous approaches rely on natural variation in the data to draw inference via multivariate models. In contrast, we exogeneously vary the prices offered to different customers using a series of randomized, split-sample field tests. This experimental design allows us to evaluate the impact of price promotions from between-group comparisons, which avoids potential problems associated with model specification, endogeneity, and/or intervening events. Second, the data enable us to distinguish prospective customers from established customers. The findings are different for these two customer segments, suggesting that if we failed to distinguish between these segments we would either observe no effect or underestimate the magnitude of the long-run effects. Finally, we focus on a durable good rather than a frequently purchased packaged good. There is reason to believe that this distinction is important, particularly with respect to forward buying. Blattberg and Neslin (1990, p. 132) and others have argued that we would expect purchase acceleration (forward buying) to be more prevalent for durable goods.

The study may be interpreted as adding generalizability to the study of promotions by extending the analysis to a new product category. We will present evidence that promotions both increase customers' deal sensitivity and attract customers who have a lower preference (willingness to pay) for the brand. These findings are consistent with evidence from the packaged-goods industry (see, for example, Mela et al. 1997, Neslin and Shoemaker 1989). We also present evidence of purchase acceleration, which has previously been demonstrated in automobile (Thompson and Noordewier 1992) and packaged goods markets (Krishna 1994).

The remainder of this paper is organized as follows. In §2, we provide an overview of the direct marketing industry and describe the design of the studies. In §3, we report findings from the three studies and in §4 we reconcile the findings across the studies. Section 5 concludes the paper.

2. Design of the Studies

In this section, we provide an overview of the three studies. Because the details vary for each study, we provide a more precise description of each study in §3. The studies reported in this paper were all conducted in mail-order catalogs distributed by a single company. The company is a medium-sized firm that sells a range of approximately 450 products targeted at well-educated older customers. For confidentiality reasons we cannot identify the name of the company.

However, we can say that the products are experience goods that have many of the same features as books or software. The products are durables; consumption occurs over time and the length of time varies by consumer (the average interpurchase period is approximately 48 weeks). Repeat purchases of an item are limited to upgrades and few, if any, customers purchase the same product twice. Over their lifetime customers may purchase many different products. The products all carry the company's own brand name, have few close substitutes, and are distributed primarily through the company's own catalogs.

The three studies were conducted at different intervals using a separate catalog for each study. In each study two versions of the respective catalog were created, a Promotion version and a Control version. Customers were randomly assigned to receive one of the two versions. The process for randomizing customers varied depending upon whether the study involved established or prospective customers. Established customers were randomized based on their unique customer account numbers, while prospective customers were randomized based on their zip + 4 codes. Where possible, we use historical data to control for individual differences between the groups. However, the absence of historical data for the prospective customers prevents a complete pre- versus post-treatment comparison.

The Promotion version offered prices that were less expensive than the Control on a subset of products, and this was the only difference between the two catalogs. The prices of the other products and all of the other catalog text were the same in both versions. In particular, the prices of the test items were presented in the Promotion version with the text: "Regularly $\$x$ Sale $\$y$." To ensure that the only difference between the two catalog versions was the actual price, the prices for these items were presented in the Control catalog with identical text: "Regularly $\$x$ Sale $\$z$." The regular price ($\$x$) was the same in both versions and the sale price was lower in the Promotion version ($\$y < \$z$).[2]

The use of the "Regularly" and "Sale" description in both the Control and Promotion versions suggests that the two conditions might more appropriately be labeled as: "Discount" and "Deep Discount." We use the "Control" and "Promotion" labels purely to improve exposition. However, this issue raises the distinction between prices and price cues. There is considerable evidence that customers are sensitive to promotional cues such as "Sale" even if prices are held constant (Inman et al. 1990; Inman and McAlister 1993; Anderson and Simester 1998, 2001). Varying both the prices and price cues would have made it difficult to distinguish the two effects.

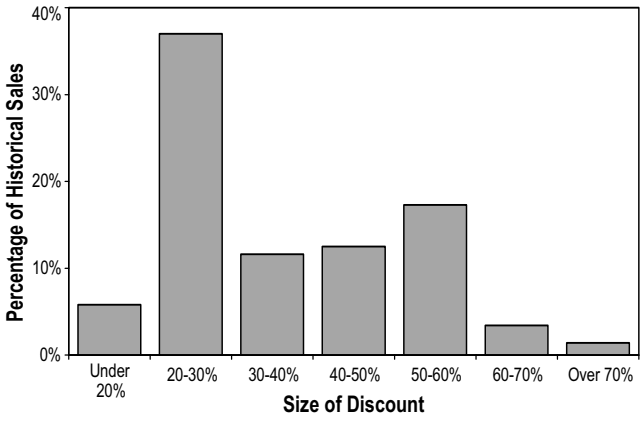
The discounts in the Control condition were on average approximately 30% lower than the regular price, while in the Promotion condition they were approximately 60% lower than the regular price. The discount level in the Promotion condition was chosen so that it would be large enough to generate an effect, but not so large that it was outside the range of historical discounts. To provide a basis for comparison, we summarize the distribution of discounts in the two years prior to the tests in Figure 1 (compiled from the company's historical sales data).

The catalog company determined which products were involved in the Test Catalog. However, the same products were used in both catalog versions, providing an explicit control for product selection. The use of a control also excludes alternative explanations arising from intervening events, such as competitive actions.

To evaluate the long-run impact of the price changes, it was important to ensure that customers received the same distribution of catalogs in the future. For this reason, the company did not distinguish between customers in the Promotion and Control conditions in catalogs mailed after the Test Catalogs (for ease of exposition we will refer to the catalogs in which the initial price tests were conducted as the "Test Catalogs"). We delay a more detailed discussion of the customer samples and catalogs used in each study until our presentation of the findings.

When customers call to place an order they are asked for the code printed on the back of the catalog from which they are ordering. This code allows the firm to identify which catalog the customer is purchasing from and, where appropriate, which catalog version (Promotion or Control). The catalog code also allows telephone operators to verify whether

Figure 1 Distribution of Historical Discounts



[2] There was an exception to this design for eight of the items in Study A. Although the price of these items was different across the two conditions ($\$y < \$x$), there was no indication that the items were discounted in either condition.

an offer has expired (every catalog explicitly states when "Sale" offers expire). The data we received contains the catalog code, an order identification number, customer identification number, quantity purchased, and price paid for each item. Items returned or cancelled by customers are netted out of the calculations.

The direct-marketing industry has a long history of conducting "split-sample" experiments, in which randomly selected customer samples are mailed modified versions of otherwise identical catalogs. Wisdom in the industry suggests that profitability depends upon a firm's ability to design, implement, and analyze such studies and then appropriately disseminate the findings within the organization. Prices are often a focus of these experiments; over 31% of catalog firms reported that they conducted split-sample experiments of pricing strategies in 1999 (Direct Marketing Association 2000). Other common experiments include testing the demand for new products, and the creative design of catalog covers, page layouts, and copy.

Testing of prices and other strategies is far more common among direct-marketing firms than among traditional retailers. In part, this reflects the benefits of conducting split-sample tests in catalogs. First, the experimental versions of the catalogs can be distributed at the same time to an identical sample of customers. Tests conducted in retail stores generally require differences in strategies over time or differences across stores. This introduces the potential for alternative explanations due to intervening events or systematic differences between stores. Second, the number and identity of catalog customers who are exposed to the different experimental versions is known. In a retail store it is much more difficult to track the total number and identity of customers who visit a store. Third, stock-outs can distort measurement of demand in a traditional retail setting. For example, there is generally no record of customers who searched for an item and then departed when they could not find it, or customers who were never aware of an item but would have purchased it if it had been on display. In a catalog setting customers initiate orders in writing or via telephone before learning whether the item is available. The decision to cancel an order, substitute an alternative item, or backorder an item is also explicit and therefore observable. Finally, catalog retailers maintain detailed records of customers' purchases, including the catalog version from which they purchased. This makes it feasible to separately target first-time and established customers. It also makes it easier to distinguish the immediate impact of a price change from the impact on future purchasing.

3. Results

In this section we present findings from the three studies. Although there are important differences in the findings, we delay discussion of these differences until §4, where we present additional analyses in an attempt to reconcile the findings across the three studies. We label the studies A, B, and C. Study A was conducted with established customers who had previously purchased from the company. Studies B and C involved prospective customers who had not previously purchased from the company. All three studies were initiated between January 1999 and April 1999, and customers' future purchases were tracked for between 22 and 28 months. Differences in the periods over which future purchasing was tracked reflect the dates that the data were extracted from the company's database.

Study A

The study was conducted in a regularly scheduled catalog containing 72 pages and 86 products. The prices charged in the two versions are summarized in Table 4. Prices on 36 of the products were approximately 40% less expensive in the Promotion version than in the Control. This was the only difference between the two catalogs; the prices of the other 50 products and all of the other catalog text were the same in both versions. The prices charged in the two versions are summarized in Table 1.

The Promotion and Control versions of the Test Catalog were distributed to separate, randomly chosen customer samples, with 18,708 customers receiving the Promotion version and 37,758 customers receiving the Control version. The company provided us with historical purchasing data for all of these customers. This data described the number of previous orders placed by each customer, the number of units ordered, the total amount spent, and the most recent order date. We also received extensive demographic data. The demographic data are based on the 1990 census data and matches customers by their

Table 1 Study A: Summary of Prices in the Promotion and Control Versions

	Control	Promotion
Prices That Varied Between Conditions		
Number of products	36	36
Average regular price	\$203.83	\$203.83
Average sale price	\$133.81	\$77.17
Prices That Did Not Vary Between Conditions		
Number of products	50	50
Average price (\$)	\$212.15	\$212.15

Notes. The price in the Promotion condition was strictly lower for all 36 products involved in the study. For the other 50 products, none of the prices differed across the two versions.

Table 2 Study A: Average of Historical Purchasing Measures for Customers That Were Mailed the Test Catalog

	Control	Promotion
Recency: Days since last order (hundreds)	6.47	6.50
Frequency: Number of orders	2.19	2.19
Monetary Value: Order amount in hundreds of dollars	1.99	1.97
Sample size	37,758	18,708

Notes. Recency and Monetary Value are both measured in hundreds. The means in the Control and Promotion are not statistically different ($p > 0.3$).

mailing address. Comparison of the historical purchasing and demographic data across the two experimental conditions confirmed that the allocation of customers was random. In Table 2 we report the average number of days since the last order (Recency); the average number of previous orders (Frequency), and the average prior-order amount (Monetary Value) for customers in each of the conditions.[3] These three measures (Recency, Frequency, and Monetary Value) are widely used within the catalog industry to segment customers when making catalog mailing decisions. There are no significant differences between conditions in these benchmarks, despite the very large sample sizes. We also considered a range of alternative measures, such as the historical average unit price, the number of days since the first order, and the various demographic measures. This analysis further confirmed the absence of systematic differences between the two customer samples.

Customers who purchased from the Test Catalog were identified and their repeat purchasing tracked for a period of 28 months. A summary of the initial and repeat purchases by these customers is presented in Table 3 and illustrated in Figure 2. As expected, the response to the Test Catalog was consistent with a downward-sloping demand function. A higher percentage of customers purchased from the Test Catalog in the Promotion condition, and on average they ordered more units. Although the experimental manipulation led to a lower average unit price in the Promotion condition, this was outweighed by the increase in units ordered, resulting in higher revenue per 1,000 customers mailed. As a basis for comparison, the industry average response rate to retail catalogs mailed in 1997 was 2.7%, with an average order size of $82 (Direct Marketing Association 1998).

We focus the remainder of our analysis on customers who purchased from the Test Catalog. This

[3] Readers may wonder how the average Recency of approximately 650 days (Table 1) can be reconciled with an average interpurchase period of 48 weeks. Recency and the interpurchase period are not directly comparable, as the Recency measure includes large values for inactive customers who made one purchase but are not expected to make another purchase. For these customers, the interpurchase time is not defined but Recency is large.

Table 3 Study A: Summary of Purchases from the Test Catalog and Future Catalogs

	Control	Promotion	Difference
Response to the Test Catalog			
Customers mailed the Test Catalog	37,758	18,708	
Customers who purchased from the Test Catalog	761	597	
Percentage that purchased	2.02	3.19	1.17**
Customers Who Purchased from the Test Catalog			
Purchases from the Test Catalog			
Number of customers	761	597	
Units ordered per customer	1.59	2.14	0.55**
Average unit price ($)	124.03	78.51	−45.52**
Repeat purchases from future catalogs			
Number of customers	761	597	
Units ordered per customer	7.67	6.89	−0.78**
Average unit price ($)	95.51	84.86	−10.65**

Notes. Significance tests are two-tailed tests of the null hypothesis that there is no difference between the Control and Promotion conditions.
*Significantly different from zero, $p < 0.05$.
**Significantly different from zero, $p < 0.01$.

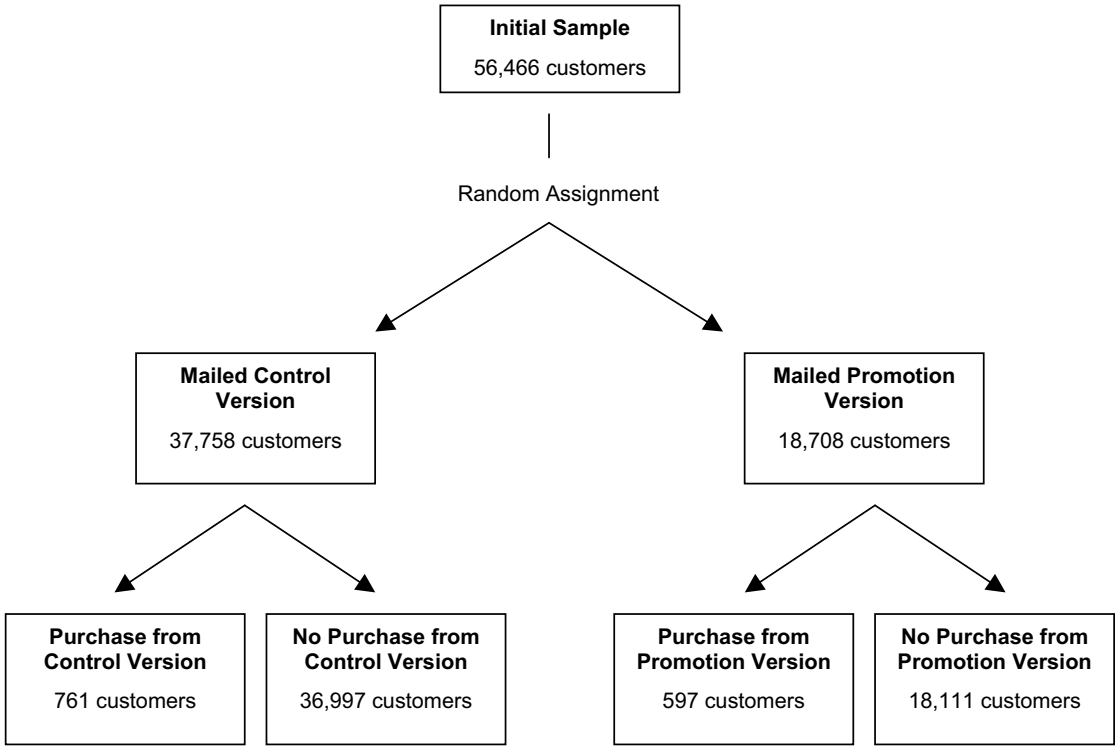
includes 597 customers in the Promotion condition and 761 customers in the Control (recall that almost twice as many customers were mailed the Control version). Analysis of repeat purchases from subsequent catalogs by these customers reveals several related findings. First, customers in the Promotion condition tended to purchase less expensive items in their future orders (the average price per unit was $84.86 versus $95.51). Recall that customers in both conditions received the same future catalogs, and so the difference cannot be explained by differences in the product offerings. Rather, the customers acquired in the two conditions chose different-priced items from the same set of product offerings. Second, on average, customers in the Promotion condition purchased fewer units from future catalogs than customers in the control (6.89 versus 7.67). Third, as a result of these two effects, overall future revenue per customer was sharply lower among customers in the Promotion condition ($584.68 versus $733.50). Finally, there is no significant difference in the total number of units purchased by these customers from both the Test Catalog and future catalogs ($1.59 + 7.67$ is not significantly different from $2.14 + 6.89$).

In separate analyses we investigated the subsequent purchasing behavior by customers who were mailed the Test Catalog but did not purchase from it. These findings are reported in Anderson et al. (2003).

Study B

There were minor differences in the catalog used for Study A and Study B, and a major difference in the customer samples. Focusing first on the customer samples, the catalog in Study B was mailed to households that had not previously purchased from the

Figure 2 Design of the Study



company. This contrasts with Study A, which was sent to customers who had previously purchased from the company. This difference allows us to investigate whether there are differences in the long-run effects for established versus new customers.

The names and addresses of the households mailed catalogs in Study B were rented from a third-party mailing vendor and were screened to have similar demographics as established customers. An extensive merge/purge process was used to compare the names and addresses of the prospective customers on the mailing list with the company's existing customers. This process ensured that none of the prospective customers that received catalogs had previously purchased from the firm.

Renting mailing lists from third-party sources is a common practice in the catalog industry. On average, two-thirds of new customers are identified from rented lists. Other sources include customer referrals and advertising in magazines and/or newspapers. Intermediaries facilitate the aggregation and rental of mailing lists and typically charge between $60 and $120 per thousand names. When choosing which lists to rent, firms typically identify the demographic characteristics of their existing customers and try to match them with the characteristics of customers on other lists. Other catalogs are the primary source of mailing lists, providing approximately 56% of the mailing lists used to identify prospective customers (Direct Marketing Association 1999, 2001). Renting mailing lists provides a reliable source of profit for many catalog firms, although approximately 50% of firms choose not to share their mailing lists with other firms.

When renting a mailing list, firms generally acquire the right to mail to these customers only once. The agreements require that the company deletes from its database all information about households that do not respond to this mailing. Mailing lists are seeded with disguised names that allow the third-party vendor to detect violations. As a result, we do not have any information about customers who did not respond to the Test Catalog. We obviously also do not have any historical purchasing data for any of these prospective customers.

Study B was conducted using an eight-page catalog that contained a total of 16 products (recall that the catalog used in Study A had 72 pages and 74 products). Among the products involved in the study, prices were on average 47% lower in the Promotion version. The two versions of the Test Catalog were distributed to separate randomly chosen customer samples, with 148,703 customers receiving the Promotion version and 148,702 customers receiving the Control version. Repeat purchasing by customers who purchased from the Test Catalog was then tracked for a period of 24 months (four months less than in Study A). Customers who purchased from the Test Catalog were mailed the same future catalogs as the customers in Study A. These future catalogs did not vary between the two conditions.

Table 4 Study B: Summary of Prices in the Promotion and Control Versions

	Control	Promotion
Prices That Varied Between Conditions		
Number of products	14	14
Average regular price ($)	207.09	207.09
Average sale price ($)	129.09	68.52
Prices That Did Not Vary Between Conditions		
Number of products	2	2
Average price ($)	69.95	69.95

Notes. The price in the Promotion condition was strictly lower for all 14 products involved in the study. For the other two products none of the prices differed across the two versions.

Purchases from the Test Catalog and future catalogs are summarized in Table 5. Initial purchasing is again consistent with a downward-sloping demand function. There were more initial purchases from the Test Catalog in the Promotion condition and on average these customers ordered more units. Consistent with the experimental manipulation, the average unit price was lower in the Promotion condition. Overall, the increase in units ordered outweighed the difference in average unit price, so that total revenue was again larger in the Promotion condition.

In the analysis of repeat purchasing we again separately consider the number of units purchased from future catalogs and the average price of these items. Focusing first on the prices, customers in the Promotion condition tended to purchase less expensive items from future catalogs, although the difference is not statistically significant. This finding is directionally consistent with the previous study. In contrast, the number of units purchased from future catalogs reveals an important difference between the two studies. In Study A, customers in the Promotion condition purchased significantly ($p < 0.01$) fewer items on average from future catalogs than those in the Control. In Study B this finding was reversed; customers in the Promotion condition purchased more items on average from future catalogs ($1.82 > 1.59$). Although this reversal is striking, we caution that the finding for Study B is only marginally significant ($p < 0.10$). To investigate whether it is robust, we replicated Study B in a third study.

Table 5 Study B: Summary of Purchases from the Test Catalog and Future Catalogs

	Control	Promotion	Difference
Response to the Test Catalog			
Customers mailed the Test Catalog	148,703	148,702	
Customers who purchased from the Test Catalog	302	560	
Percentage that purchased	0.20	0.38	0.18**
Customers Who Purchased from the Test Catalog			
Purchases from the Test Catalog			
Number of customers	302	560	
Units ordered per customer	1.35	1.56	0.21**
Average unit price ($)	114.02	73.99	−40.03**
Repeat Purchases from Future Catalogs			
Number of customers	302	560	
Units ordered per customer	1.59	1.82	0.23
Average unit price ($)	95.94	92.30	−3.64

Notes. Significance tests are two-tailed tests of the null hypothesis that there is no difference between the Control and Promotion conditions.

*Significantly different from zero, $p < 0.05$.

**Significantly different from zero, $p < 0.01$.

Study C

The design of Study C was almost identical to Study B and focused only on households that had not previously purchased from the company. This study was conducted using a 16-page catalog that included a total of 36 items. The Promotion version offered deeper discounts on 32 of these items. The two versions were in all other respects identical, as were the future catalogs mailed to customers who purchased from either version. Repeat purchasing was tracked for a total of 22 months. The price differences between the two versions are summarized in Table 6 and the initial and repeat purchases are described in Table 7. We again have no information about customers who did not purchase from either version of the Test Catalog and no historical purchasing data for any of these prospective customers.

The response to the Test Catalog is similar to that in Studies A and B. Demand is downward sloping, with more revenue earned per customer mailed from the Promotion version of the Test Catalog. More importantly, the two future-purchasing findings in Study B survived replication in Study C. First, the average price of the units ordered from future catalogs was significantly ($p < 0.05$) lower in the Promotion condition than in the Control. Second, the number of units purchased from future catalogs was higher in the Promotion condition, with customers ordering 1.33 units on average in the Promotion condition and

Table 6 Study C: Summary of Prices in the Promotion and Control Versions

	Control	Promotion
Prices That Varied Between Conditions		
Number of products	32	32
Average regular price ($)	210.89	210.89
Average sale price ($)	134.01	77.14
Prices That Did Not Vary Between Conditions		
Number of products	4	4
Average price ($)	174.95	174.95

Notes. The price in the Promotion condition was strictly lower for all 32 products involved in the study. For the other four products, none of the prices differed across the two versions.

Table 7 Study C: Summary of Purchases from the Test Catalog and Future Catalogs

	Control	Promotion	Difference
Response to the Test Catalog			
Customers mailed the Test Catalog	97,847	146,774	
Customers who purchased from the Test Catalog	225	588	
Percentage that purchased	0.23	0.40	0.17**
Customers who Purchased from the Test Catalog			
Purchases from the Test Catalog			
Number of customers	225	588	
Units ordered per customer	1.49	1.94	0.45**
Average unit price ($)	115.41	81.68	−33.73**
Repeat Purchases from Future Catalogs			
Number of customers	225	588	
Units ordered per customer	0.99	1.33	0.34**
Average unit price ($)	106.36	95.49	−10.87*

Notes. Significance tests are two-tailed tests of the null hypothesis that there is no difference between the Control and Promotion conditions.
*Significantly different from zero, $p < 0.05$.
**Significantly different from zero, $p < 0.01$.

just 0.99 units on average in the Control. This difference is statistically significant ($p < 0.01$) and again reverses the finding in Study A.

4. Reconciling the Three Studies

In this section we compare the findings across the three studies. To facilitate this discussion, we summarize the results from the Test Catalog and future catalogs in Table 8. For ease of comparison we report demand in the Promotion condition indexed against demand in the respective Control.

In all three studies, aggregate demand in the Test Catalog is downward sloping, with more revenue earned per customer mailed from the Promotion version of the Test Catalog. The increase in demand is reflected in both the number of customers that responded to the Test Catalog and in the number of units that they purchased.

The analysis of purchases from future catalogs focused on the number of items purchased and the average price of those items. The price analysis revealed that customers in the Promotion condition tended to purchase less expensive items from future catalogs. This result held in all three studies. We also observed a difference in the number of units purchased, although this finding varied across the studies. In Study A, customers in the Promotion condition purchased fewer units from future catalogs than customers in the Control condition. In Studies B and C the direction of the result was reversed, with customers in the Promotion condition purchasing more units from future catalogs than customers in the Control.

Table 8 The Findings in the Three Studies: Demand in the Promotion Condition Indexed to 100 in the Control

	Study A	Study B	Study C
Purchases from the Test Catalog			
Percentage that purchased	158	185	174
Units ordered per customer	135	116	130
Average unit price ($)	63	65	71
Repeat Purchases from Future Catalogs			
Units ordered per customer	90	114	134
Average unit price ($)	89	96	90

Note. The measures are all indexed to 100 in the respective Control condition: 100 ∗ Promotion/Control.

In the remainder of the paper we focus on purchases from future catalogs and investigate four explanations for the findings: forward buying (purchase acceleration), customer selection, learning, and increased deal sensitivity. We conclude the section with an integrated model that allows us to partial out the marginal effect of each explanation.

Forward Buying

In all three studies, customers in the Promotion condition ordered more units on average from the Test Catalog. This may reflect forward buying, in which customers take advantage of the low prices to stock up on products required for future consumption (Krishna 1992, 1994; Thompson and Noordewier 1992). This interpretation is consistent with the nature of the products. They share similar characteristics with books: They are not perishable and customers will typically consume them only once, so that further consumption requires an additional purchase. It is also consistent with the finding in Study A that there is no significant difference in the total number of units purchased from the Test and future catalogs ($1.59 + 7.67$ is not significantly different from $2.14 + 6.89$; see Table 3).

We would expect forward buying to result in three effects: (a) higher demand in the Promotion condition immediately after the Test Catalog was mailed, (b) a subsequent drop-off in demand while customers consume their inventory, and (c) demand eventually returning to the same level in both conditions. Given that the average interpurchase time is less than a year (48 weeks), we expect forward buying to have its strongest impact on demand in the first 12 months after the Test Catalog. Therefore, if we focus only on downstream purchases that occur more than 12 months after the Test Catalog, we can mitigate the effects of forward buying. In Table 9 we separately report the number of units ordered within the first 12 months of the Test Catalog and units ordered more than 12 months after the Test Catalog. For ease of exposition, we report demand in the Promotion condition indexed against demand in the respective Control.

Table 9 Units Ordered per Customer from Future Catalogs

	Study A	Study B	Study C
Purchases within 12 Months of the Test Catalog	87	111	123
Purchases more than 12 Months after the Test Catalog	92	120	149

Note. The measures are all indexed to 100 in the respective Control condition: 100 ∗ Promotion/Control.

The findings in Table 9 offer some support for the forward-buying explanation. In all three studies, demand in the Promotion condition (relative to the Control) was lower in the first 12 months after the Test Catalog was mailed compared to demand in subsequent months ($87 < 92$, $111 < 120$, $123 < 149$). This pattern of findings does not depend on the 12-month demarcation; it holds when varying the length of the first period.

We conclude that there is some evidence supporting the forward-buying explanation; the reduction in (relative) demand in the Promotion condition immediately after the Test Catalog is consistent with customers consuming the inventory from their Test Catalog purchases. However, even if we only consider purchases made more than 12 months after the Test Catalog was mailed, the deep discounts in the Promotion condition are associated with a decrease in repeat purchases in Study A ($92 < 100$) and an increase in repeat purchasing in Studies B and C ($120 > 100$ and $149 > 100$). It seems that forward buying cannot provide a complete explanation for these findings.

It is not obvious how forward buying could explain the tendency for customers in the Promotion condition to purchase less expensive items from future catalogs. However, we repeated the analysis in Table 9 with the price data. Perhaps not surprisingly, the price effect did not vary systematically over time.

Selection

Our analysis of future purchasing only considered customers who purchased from the Test Catalog. Because prices were lower in the Promotion version, customers with low valuations for the company's products may have been more likely to purchase in that condition. This raises the possibility of "selection," in which purchasers in the Promotion sample on average have a lower valuation for the company's products than do those in the Control sample (Neslin and Shoemaker 1989). This selection effect could account for fewer repeat purchases and a preference for less expensive items in the Promotion sample.

The historical purchasing data in Study A offers an opportunity to investigate this selection explanation. In Table 10 we report the average of the historical Recency, Frequency, and Monetary Value (RFM) measures for customers who *purchased* from the Test Catalog in Study A. These averages confirm that there are systematic differences in the customer samples. Customers in the Promotion condition on average purchased less recently ($p < 0.05$), had fewer prior orders ($p < 0.01$), and had a lower average order amount ($p < 0.10$).[4]

Table 10 Study A: Average of Historical and Demographic Variables for Customers Who Purchased from the Test Catalog

	Control	Promotion	Difference
Recency: Days since last order	2.86	3.27	0.41*
Frequency: Number of orders	5.58	4.66**	−0.92**
Monetary value: Average order amount	2.00	1.88	−0.12
Sample size	761	597	

Notes. Recency is measured in hundreds of days and Monetary Value is measured in hundreds of dollars. Significance describes the probability level at which we can reject the null hypothesis that there is no significant difference in the averages between the Control and Promotion conditions.

*Significantly different from zero, $p < 0.05$.

**Significantly different from zero, $p < 0.01$.

The differences in prior purchasing patterns offer strong support for the selection argument. Given that the Promotion condition attracted customers who historically ordered smaller amounts and ordered less frequently, we would expect customers in this condition to purchase less expensive items and purchase fewer items from future catalogs. To further investigate this explanation, we used multivariate analysis to explicitly control for customers' past purchasing patterns.

The multivariate analysis required separate models to investigate the number of items and average price results. In particular, the number of units purchased is a count measure, which might be expected to follow a Poisson distribution. Because there are customers with zero purchases, it might appear that the number of orders received from future catalogs is censored. However, when we recognize that the measure is a count measure, it is clear that values of zero are naturally occurring and not a result of any censoring or truncation. The Poisson specification is ideally suited to estimating count models and naturally accommodates observations of zero. In contrast, the Tobit model, which adjusts for censoring, is not appropriate both because it assumes censoring and because it requires a continuous dependent variable.

By implementing a Poisson model, we assume that the number of units ordered by customer i from

[4] We found few differences in the demographic variables. This may reflect the imprecision in the demographic data; they are not specific measures for each household, and instead reflect average measures for the census block in which the household is located. In this study, the median number of households in a census block is 497.

future catalogs (Q_i) is drawn from a Poisson distribution with parameter λ_i:

$$\text{Prob}(Q_i = q) = \frac{e^{-\lambda_i}\lambda_i^q}{q!}, \qquad q = 0, 1, 2, \ldots, \quad (1)$$

where: $\ln(\lambda_i) = \mathbf{\beta X}_i$. The historical RFM measures are well-established metrics for segmenting customers in this industry and so provide natural candidates for explanatory variables. We also include a dummy variable identifying whether customer i received the Promotion version of the Test Catalog:

$$\mathbf{\beta X}_i = \beta_0 + \beta_1 Recency_i + \beta_2 Frequency_i + \beta_3 Monetary\ Value_i + \beta_4 Promotion_i. \quad (2)$$

We do not include variables describing the characteristics of the future catalogs (including the price of the items in those catalogs), as the catalogs were the same for all customers.

Heckman (1979) proposed an alternative approach for addressing selection bias. This approach uses two equations: a primary response equation and a sample-selection equation describing when a response is observed. To implement the Heckman approach we included the RFM measures in the first-step Probit model and the RFM measures in the second-step Poisson response model. This restricts the impact of the RFM measures to the indirect influence of selection. In the second stage, selection is captured in a single variable referred to as the Mills ratio, which is a nonlinear transformation of the RFM variables. Equation (2) controls for selection in a more direct, flexible manner: The RFM measures enter as three separate variables, providing controls for selection and the direct influence of these measures on the number of units purchased. Although the two approaches yield almost identical findings, Equation (2) offers the additional advantages of simplicity and ease of interpretation.

To investigate the average price of items purchased from future catalogs (*Future Price*$_i$), we used multivariate regression (OLS). The historical RFM and *Promotion* variables again provide natural candidates for explanatory variables. To complement these measures, we also included a more direct historical control for this model: the average price of items in prior orders (*Historical Price*$_i$). This led to the following specification:

$$Future\ Price_i = \beta_0 + \beta_1 Recency_i + \beta_2 Frequency_i + \beta_3 Monetary\ Value_i + \beta_4 Historical\ Price_i + \beta_5 Promotion_i + \varepsilon. \quad (3)$$

In the Poisson count model we included all of the customers who purchased from the Test Catalog, while in the price model we only considered customers who purchased at least one item from a future catalog. This restriction to customers who made at least one subsequent purchase suggests that it may be appropriate to adjust the coefficients to account for the effects of selection. However, we are only interested in how the deep discounts affected the *Average Unit Price* by those customers who purchased again, which is measured by the unadjusted coefficients. Adjusting the coefficients to account for selection would take into account how the promotion would have affected the price paid by customers who did not make another purchase. *Because* these customers did not purchase again, we are not concerned with what price they would have hypothetically paid.

We report the findings when estimating both models using all of the downstream data in Table 11 (we later discuss how the *Number of Units* model is affected when we restrict the dataset to control for forward buying). The findings indicate a strong association between historical purchasing behavior and purchases from future catalogs. Customers who historically purchased more frequently, more recently, and who spent more on each purchase also tended to purchase more units from future catalogs. In the price model, there was a very strong relationship between the price of items purchased from future catalogs and both the *Monetary Value* and *Historical Price* variables.

The coefficients for the *Promotion* variable in the quantity and price models help to reveal whether selection is a complete explanation for our earlier findings. The *Promotion* variable is not significant in the quantity model. After controlling for the historical RFM measures, there is no longer a significant relationship between the version of the Test Catalog that customers received and the number of units purchased from future catalogs. It appears that selection can explain the earlier finding (in Study A) that deeper discounts lead to fewer subsequent purchases. In the *Average Unit Price* model the *Promotion* variable

Table 11 Controlling for Selection

	Units per Customer		Average Unit Price	
Intercept	1.803**	(0.022)	68.642**	(4.000)
Recency	−0.117**	(0.004)	0.168	(0.462)
Frequency	0.043**	(0.001)	0.538**	(0.228)
Monetary Value	0.093**	(0.005)	6.316**	(1.297)
Historical Price			13.685**	(2.436)
Promotion	−0.014	(0.020)	−10.749**	(2.825)
Model	Poisson		OLS	
Adjusted R^2			0.12	
Log likelihood	−7,066			
Null log likelihood	−8,487			
Sample size	1,358		1,081	

Notes. Standard errors are in parentheses.
*Significantly different from zero, $p < 0.05$.
**Significantly different from zero, $p < 0.01$.

is significant ($p < 0.01$). This suggests that even after controlling for selection, customers in the Promotion condition were more likely to purchase less expensive items from future catalogs.

Unfortunately, the absence of historical purchasing data for first-time customers limits this analysis to Study A. While we cannot perform a similar analysis for Studies B and C, we note that it is unlikely that selection offers a complete explanation. It is difficult to identify a selection effect such that the Promotion condition attracted customers who tended to purchase lower-priced items and purchase *more* items from future catalogs. Moreover, a complete explanation would need to address the reversal of the future demand findings between Study A and Studies B and C.

We conclude that there remain two unexplained findings. First, we do not yet have an explanation for why future demand was higher in the Promotion condition in Studies B and C, or why this reverses the findings in Study A. Second, selection and forward buying do not completely explain the tendency for customers in the Promotion condition to purchase less expensive items from future catalogs. We will present separate explanations for these two findings. We first consider whether customer learning may help to explain the variation in the quantity purchased from future catalogs and later investigate whether changes in deal sensitivity may help to explain the differences in the average unit price.

Learning

Evidence from a variety of sources suggests that promotions may contribute to customer learning. Customers are often uncertain about the quality of different product features and/or the price that will be charged on future purchase occasions. Offering promotions to these customers may lead to more favorable price and/or quality expectations. This link between promotions and favorable expectations may be either direct or indirect. Erdem and Keane (1996) offer an example of an indirect link, arguing that promotions increase consumption, which in turn may lead to favorable quality perceptions. A more direct link is drawn in several signaling models, which predict that low initial prices signal favorable information about quality (Schmalensee 1978, Dawar and Sarvary 1997) or future price levels (Bagwell 1987, Simester 1995).

The learning explanations may also help to explain the difference in findings. Studies B and C were conducted with first-time customers who had not purchased from the catalog before and generally had not seen previous catalogs. As a result, they had relatively little information about the company and its products. In contrast, the established customers in Study A had previously purchased from the company and could use information from previous purchases and catalogs to form inferences about the prices and quality of products in future catalogs. For this reason, we would expect the first-time customers, who had almost no other information with which to form price and quality expectations, to be more sensitive to any learning effects.

Although the "established" customers in Study A had all made a prior purchase from the company, the frequency and recency of their prior purchases varies. This suggests an opportunity to further investigate the learning explanations. Customers with more *frequent* prior purchases, and customers whose prior purchases were more *recent*, will tend to have more alternative sources of information with which to form price and quality expectations. We might anticipate that the expectations of these customers would be less sensitive to the prices in the Test Catalog. To investigate this prediction, we grouped the customers in Study A based on the recency and frequency of their prior purchases. In particular, we used the median number of prior purchases (three prior purchases) and the median number of days since the last purchase (154 days) to assign the customers to one of four groups. We then compared future demand in the Promotion and Control conditions in each of the four groups. The findings are reported in Table 12, where we report the average number of units ordered from future catalogs.

The learning explanations predict that low prices in the Promotion condition will lead to favorable expectations. Customers with fewer prior purchases and less-recent prior purchases have fewer alternative sources of information and so will be more sensitive to this effect. The findings support this claim.

Table 12 Study A: Average Number of Units Ordered from Future Catalogs by Frequency and Recency of Customers' Prior Purchases

	Historical Order Frequency and Recency			
	Infrequent + Not Recent	Infrequent + Recent	Frequent + Not Recent	Frequent + Recent
Units per Customer				
Control	2.45	6.75	7.26	11.93
Promotion	3.18	6.44	6.72	11.17
Difference	0.73	−0.31	−0.54	−0.76
Index†	130	95	93	94
Sample Size				
Control	182	130	180	269
Promotion	186	103	131	177

Notes. †*Index* = 100 ∗ Promotion/Control. The table only includes customers who purchased from the Test Catalog. "Infrequent" prior purchases denotes customers with one or two prior purchases. "Recent" prior purchases denotes customers who purchased within 153 days of the mail date for the Test Catalog "Frequent" and "Not Recent" denote customers with more prior purchases or less-recent prior purchases, respectively.

The adverse effect of the Promotion on future demand was limited to customers who had a history of recent and/or frequent prior purchases. Among customers who had not purchased recently and had few prior purchases, the Promotion condition was associated with an *increase* in future demand (130 > 100). The results for these customers mirror the results in Studies B and C.[5]

The findings in Table 12 do not control for selection and forward buying. However, the analysis in this table does suggest an approach for distinguishing between these alternative explanations. In particular, the findings show that the impact of the Promotion on the number of units purchased from future catalogs is moderated by the recency and frequency of customers' historical purchases. We can explicitly investigate this hypothesis by adding dummy variables to Equation (2) identifying each of these four customer segments:

$$\begin{aligned}\boldsymbol{\beta}\mathbf{X}_i = &\sum_{j=1}^{4}\alpha_j Segment\ j_i + \beta_1 Recency_i + \beta_2 Frequency_i \\ &+ \beta_3 Monetary\ Value_i \\ &+ \sum_{j=1}^{4}\tau_j Segment\ j_i * Promotion_i. \qquad (2a)\end{aligned}$$

Under this specification the τ_j coefficients identify the Promotion effect on the respective segments (the use of dummy variables allows for nonlinear relationships). We report the results of this analysis in Table 13, where we focus on purchases made at least 12 months after the Test Catalog (including all downstream purchases yielded a very similar pattern of results). For the sake of brevity we omit the four constants. The findings confirm that among customers who had not purchased recently and had few prior purchases, the Promotion condition was associated with an *increase* in future demand.

Evidence that the difference in prices between the two Test Catalog versions may have affected customer learning begs the question: What did customers learn? Unfortunately, the data offer little help in answering this question. However, a review of the literature suggests at least two alternatives. In offering these two alternatives we recognize both that other information explanations may exist, and that the explanations are not mutually exclusive, so that findings may reflect a combination of these alternatives.

First, the prices in the Test Catalog may have provided information about the quality of the products.

Table 13 Poisson Regression of Units Purchased 12 Months After the Test Catalog

Variables	Coefficients
Recency	−0.054** (0.008)
Frequency	0.036** (0.002)
Monetary Value	0.098** (0.008)
Promotion* (infrequent + not recent)	0.262** (0.081)
Promotion* (frequent + not recent)	0.017 (0.057)
Promotion* (infrequent + recent)	−0.125 (0.071)
Promotion* (frequent + recent)	0.027 (0.037)
Log likelihood	−5,207
Null log likelihood	−6,136
Sample size	1,358

Notes. Standard errors are in parentheses.
*Significantly different from zero, $p < 0.05$.
**Significantly different from zero, $p < 0.01$.

There is a well-established literature arguing that low introductory prices may signal favorable information about product quality (Schmalensee 1978, Milgrom and Roberts 1986, Tirole 1988, Dawar and Sarvary 1997). The argument recognizes that high-quality firms can expect high repeat-purchase rates, and so are more willing to offer initial discounts to induce trial.

Second, the prices in the Test Catalog may have provided information about the prices of products in future catalogs. The extent to which customers search the company's catalogs on future purchase occasions may be influenced by their expectations about future prices. If low prices in the Promotion condition led to expectations that future prices would also be lower (compared to the Control), this could explain the increase in future demand. There is a growing literature examining the information role played by prices. For example, Bagwell (1987) shows that a low introductory price may serve as a credible signal of a firm's cost type when consumers are uncertain about the identity of the efficient firm. Consumers that make repeat purchases rationally expect that a firm that offers low introductory prices will also offer lower prices than the competition in a future period. Other related models include Milgrom and Roberts (1986), Bagwell and Riordan (1991), Simester (1995), and Anderson and Simester (1998).

Interestingly, while these two learning explanations are consistent with the findings in this study, they conflict with evidence in the behavioral literature that customers either attribute low prices to poor product quality (object perception) or attribute their purchasing of a promoted product to a weak preference for the product (self-perception).[6] Both of these attribution theories predict that deep discounts in the Promotion

[5] They also suggest that the difference in the findings across studies is at least in part due to customer differences, rather than simply differences between the catalogs used in the three studies (see later discussion).

[6] See Blattberg and Neslin (1990, Ch. 2) for a more comprehensive discussion of attribution and how to apply it to the study of price promotions.

condition would have led to a reduction in future demand.

Behavioral learning theory may also offer an alternative explanation. A central prediction from this theory is that rewarded behavior will often persist even when rewards are withdrawn (Rothschild and Gaidis 1981, Blattberg and Neslin 1990). The deep discounts may have been more effective in training customers to purchase when they see a discount cue.[7] This explanation represents a psychological rationale for customer learning, which has identical implications and is therefore difficult to distinguish from the economic rationale offered by the signaling arguments.

We conclude that the findings from the three studies provide evidence to both support and limit the application of customer learning arguments. While the findings in Studies B and C offer support, the findings in Study A suggest that learning does not explain the behavior of all customers. In Study A, low prices in the Promotion condition did not appear to prompt favorable expectations among customers who had recent and/or frequent prior purchase experiences.

Price Sensitivity

We have not yet offered a complete explanation for why customers in the Promotion condition purchased less-expensive items from future catalogs. In this section we offer an explanation that is motivated by the evidence in the preceding section that the Test Catalog affected the future demand of different customer segments in different ways. In this section we investigate whether the effect of the Test Catalog on the price of the items purchased from future catalogs also differed across customer segments. In particular, we grouped customers into four segments of approximately equal size based on the average *Historical Price* of their prior purchases. We then compared the average price of items purchased from future catalogs in each of the four segments. The findings are reported in Table 14.

The findings confirm that the price of the items in customers' historical orders moderated the relationship between the Promotion condition and the average price of items ordered downstream. The difference between the Promotion and Control conditions was larger among customers for whom the *Historical Price* of prior orders was above average (Segments C and D). We caution that this comparison does not control for the effects of selection. To investigate whether the results in Table 14 survive when we control for selection, we added dummy variables identifying the four segments to Equation (3):

$$\begin{aligned} Future\ Price_i &= \sum_{j=1}^{4} \alpha_j Segment\ j_i + \beta_1 Recency_i + \beta_2 Frequency_i \\ &\quad + \beta_3 Monetary\ Value_i + \beta_4 Historical\ Price_i \\ &\quad + \sum_{j=1}^{4} \tau_j Segment\ j_i * Promotion_i + \varepsilon. \end{aligned} \tag{3a}$$

The τ_j coefficients identify the Promotion effect on the respective segments (and allow for nonlinear relationships). We report the results of this analysis in Table 15 (omitting the four constants). The findings confirm that the negative association between the Promotion condition and the price of items purchased downstream is much stronger for customers who had paid higher prices in the past. For completeness, we also evaluated a specification that included interactions between the *Recency* and *Frequency* measures and the *Promotion* variable. The coefficients for these interactions were not significant and inclusion of these interactions did not change the pattern of findings.

The marketing literature offers several possible interpretations for this finding. They include a recent

Table 14 Study A: Average Unit Price of Orders from Future Catalogs by Monetary Value of Customers' Prior Orders

	Historical Price Segment			
	A (Lowest)	B	C	D (Highest)
Average Unit Price				
Control ($)	78.75	95.09	111.67	124.53
Promotion ($)	75.66	86.96	95.07	111.52
Difference ($)	3.09	8.13	16.54	13.01
Index†	96	91	85	90
Sample Size				
Control	132	153	170	159
Promotion	139	116	100	112

Notes. †*Index* = 100 ∗ Promotion/Control. The table only includes customers who purchased from the Test Catalog. The cutoffs for the four segments occurred at *Historical Price* levels of $88, $122, and $162.

Table 15 Regression of Average Unit Price

Variables	Coefficients
Recency	−0.229 (0.461)
Frequency	0.405 (0.240)
Monetary Value	5.682** (1.312)
Historical Price	−2.191** (0.442)
Promotion* Segment A	−3.094 (5.537)
Promotion* Segment B	−7.448 (5.612)
Promotion* Segment C	−15.808** (5.756)
Promotion* Segment D	−14.639** (5.650)
Adjusted R^2	0.13
Sample size	1,081

Notes. Standard errors are in parentheses.
*Significantly different from zero, $p < 0.05$.
**Significantly different from zero, $p < 0.01$.

[7] The authors are grateful to the area editor for proposing this alternative explanation.

series of papers showing that category promotions may increase customers' price or deal sensitivities (Mela et al. 1997, 1998; Jedidi et al. 1999). These studies all involved repeated promotions, whereas the experimental manipulations in this paper were confined to a single promotion. However, it is possible that the deep discounts in the Promotion condition were sufficient to make customers in this condition more reluctant to purchase at higher prices in the future. We would expect any such shift in price sensitivity to be particularly large for customers who had paid higher prices in the past (and therefore revealed themselves to be less price sensitive in the past). A related group of theories argues that customers evaluate a transaction against a reference price (Thaler 1985, Kalyanaram and Winer 1995). If deep discounts in the Promotion condition lowered customers' reference prices, these customers will be more likely to seek out lower-priced items in the future. We would again expect a bigger effect among customers who had paid higher prices in the past (revealing a higher historical reference price). These two theories are complementary; the reference price argument may provide an explanation for the shift in price and/or deal sensitivities. Notably, both explanations predict a change in customers' underlying preferences, and so differ from the selection, forward-buying, and learning explanations.

To further distinguish between price and deal sensitivity, we examined how many items were purchased at a promotional (sale) price both before and after the Test Catalog. As expected, the randomized experimental design ensured that there were no differences between the Promotion and Control groups in the historical proportion of items purchased on sale. However, following the Test Catalog, customers in the Promotion condition were more likely to purchase at a promotional price than customers in the Control. This difference is not statistically significant for customers who historically paid the lowest average price ($p > 0.10$, Segment A), but is statistically significant for the other segments ($p < 0.05$, Segments B, C, D). Further investigation revealed that customers who had historically paid the highest prices (Segment D) not only became more likely to seek out promoted items, they also became more likely to seek out *lower-priced* promoted items.

We conclude that the tendency for customers in the Promotion condition to purchase less-expensive items downstream appears to be explained at least in part by increased price and/or deal sensitivity. This effect is particularly strong among customers who paid higher prices in the past.

Other Factors

Several differences between the three studies may also have contributed to the variation in results across the three studies. We summarize these differences in Table 16. The differences reflect the practical reality of conducting field studies of this size.

Table 16 Differences in the Study Designs

	Study A	Study B	Study C
Average % discount in Promotion version	42	47	42
Number of prices varied	36	14	32
Number of pages	72	8	16
Number of products	86	16	36
Number of months of future data	28	24	22

The catalog used in Study A contained considerably more products and pages than the catalogs used in Studies B and C, but we have two reasons to believe that these catalog differences were not critical. First, within Study A we found effects similar to Studies B and C (see Tables 12 and 13). Second, there was a small sample of additional customers who received the Test Catalog used in Study A. These customers had not previously purchased from the company, but had called and requested a catalog. In the direct-mail industry these contacts are commonly referred to as "Inquirers." The Control and Promotion versions of the Test Catalog used in Study A were mailed to 5,829 and 2,932 Inquirers, respectively (these customers were not included in our earlier analyses). A total of 47 (0.81%) and 46 (1.57%) Inquirers purchased from the Test Catalog in the two conditions, respectively. Their long-run response was analogous to Studies B and C. Customers in the Promotion condition purchased more units on average from future catalogs (3.30 versus 3.00) and chose lower-priced items ($74.60 versus $99.70). Given that these Inquirers received the same version of the Test Catalog as other customers in Study A, the difference in their long-run response cannot be attributed to catalog differences.

We also considered two other alternative explanations. First, the difference in the number of months of future data varied across the studies. However, we replicated all of our univariate analyses using 22 months of future data for all three studies and observed the same pattern of findings. Second, customers in the Promotion condition for Study A purchased more items from the Test Catalog. Readers may wonder whether the absence of any other popular items left to purchase might explain the reduction in demand downstream in that study. This ceiling effect requires that the catalog have a limited selection of popular items. However, the firm in this study has a broad product line of over 450 items. Historical demand among these products is approximately consistent with an 80/20 rule, with 80% of unit volume attributable to 17% of the products (or approximately 80 items). Even for high-volume purchasers

there are still a large number of popular items available to choose from. Moreover, the difference in the average number of units ordered from the Test Catalog between customers in the two conditions is less than one unit. It seems unlikely in this context that this difference is large enough for a ceiling effect to only affect customers in the Promotion sample.

Summary

The findings reported in this section show evidence of several long-run effects. To evaluate the relative importance of these effects, we estimated a series of models in which we added and/or removed controls for each of the effects. A complete specification of the models and a summary of the results is in the Appendix (see Table A.1). To measure the marginal effect of each explanation, we compared the *Promotion* coefficients across these models. The results are summarized in Table 17.

Comparison of the *Number of Units* models reveals the marginal effects of forward buying, selection, and learning. In the absence of any controls, the deep discounts in the Promotion condition are associated with a 10.8% reduction in future units ordered. However, controlling for forward buying alone mitigates this to 8.3%. We conclude that the effect due to forward buying is the difference between 10.8% and 8.3%, or approximately 2.5% (this difference is not statistically significant).[8] Controlling for selection reduces the *Promotion* effect to −1.4, suggesting that selection contributes approximately 9% to the overall effect (the difference between 10.8% and 1.4%). Among customers without recent or frequent purchases, we find an increase in demand of 26.2%, which we attribute to learning.[9]

In the *Average Unit Price* models, customers in the Promotion condition purchase items from future catalogs that are $12.55 less expensive. This difference is reduced to $10.75 after controlling for selection. Because the average price per unit is $95 (see Table 3), selection results in a 1.9% reduction in average price paid (the difference between $12.55 and $10.75 as a proportion of $95). In comparison, the increased deal sensitivity among customers who historically purchased higher-priced items (Segments C and D) led to an effect of approximately 15%.

Table 17 Summary of Promotion Effect Sizes on Number of Units and Average Unit Price

	Approximate Effect Size	
	Number of Units (%)	Average Unit Price (%)
Forward buying	3	
Selection	9	2
Learning[a]	26	
Deal sensitivity[b]		15

[a]Among customers without recent or frequent prior purchases.
[b]Among customers who purchased high-priced items prior to the test.

5. Conclusion

We have presented evidence from three large-scale field studies investigating whether promotion depth has a long-run effect on demand. Overall, the results show evidence of several long-run effects: forward buying, selection, customer learning, and increased deal sensitivity. Short-run metrics that ignore these effects overstate the overall change in demand for established customers. The implication is that if prices are set based on short-run elasticity, then they will be too low. Among first-time customers, the short-run metrics underestimate the total increase in demand. If prices are set based on short-run elasticity, then they will be too high.

The evidence that deeper promotions increase repeat-purchase rates among first-time customers is surprising. However, our replication of the effect provides reassuring evidence that the finding is robust. We have argued that the finding is consistent with customer learning. Prospective customers have little information about a firm, and a low initial price may lead to favorable expectations about future price and/or quality levels.

The three studies all investigate the long-run effects of a temporary price change. We cannot say how customers would have responded to permanent price changes, or how they might have responded to a subsequent discount. Investigating these issues would require different studies in which the experimental manipulations were maintained for a longer period, or repeated in a subsequent catalog.

Acknowledgments

The authors wish to thank Carl Mela and Birger Wernerfelt, together with seminar participants at the Universities of Chicago; Illinois; Minnesota; and Toronto; the London Business School; and USC, for their comments. They would also like to thank the company that provided the data reported in the paper. The first author thanks the Kilts Center for Marketing at the University of Chicago for research support.

Appendix: Specifications of the Models in Table A.1

Number of Units Models

1. No Controls: This model was estimated using all of the downstream data. Equation (2) was modified as follows:

$$\boldsymbol{\beta}\mathbf{X}_i = \beta_0 + \beta_4 Promotion_i. \quad (2b)$$

[8] We obtain a similar effect size by comparing a model with selection to a model with both selection and forward buying: $-1.4\% - 1.3\% = -2.7\%$.

[9] The learning and deal sensitivity effects are computed by comparing the respective promotion coefficients with a benchmark of zero.

Table A.1. Comparison of Promotion Coefficients Across Models

Number of Units Models	No Controls	Forward Buying (FB)	Selection (SEL)	SEL + FB	SEL + FB + Learning
Promotion	−0.108** (0.020)	−0.083** (0.027)	−0.014 (0.020)	0.013 (0.027)	
Promotion* (infrequent + not recent)					0.262** (0.081)
Promotion* (frequent + not recent)					0.017 (0.057)
Promotion* (infrequent + recent)					−0.125 (0.071)
Promotion* (frequent + recent)					0.027 (0.037)

Average Unit Price Models	No Controls	Selection (SEL)	SEL + Price Sensitivity
Promotion	−12.552** (2.969)	−10.749** (2.825)	
Promotion* Segment A			−3.094 (5.537)
Promotion* Segment B			−7.448 (5.612)
Promotion* Segment C			−15.808** (5.756)
Promotion* Segment D			−14.639** (5.650)

*Significantly different from zero, $p < 0.05$.
**Significantly different from zero, $p < 0.01$.

2. Forward Buying: Equation (2a) was estimated using transactions that occurred more than 12 months after the date that the Test Catalog was mailed.

3. Selection: Equation (2) was estimated using all of the downstream data. The results are also reported in Table 11.

4. Selection + Forward Buying: This model was estimated using transactions that occurred more than 12 months after the date that the Test Catalog was mailed. The model estimated is specified in Equation (2).

5. Selection + Forward Buying + Learning: Equation (2a) was estimated using transactions that occurred more than 12 months after the date that the Test Catalog was mailed. The results are also reported in Table 13.

Average Unit Price Models

These models were all estimated using all of the downstream data.

1. No Controls: Equation (3) was modified as follows:

$$Future\ Price_i = \beta_0 + \beta_5 Promotion_i + \varepsilon. \quad (3b)$$

2. Selection: This is the same model that was reported in Table 11.

3. Selection + Price Sensitivity: This is the same model that was reported in Table 15.

References

Ailawadi, Kusum, Scott Neslin. 1998. The effect of promotion on consumption: Buying more and consuming it faster. *J. Marketing Res.* **35**(August) 390–398.

Anderson, Eric, Duncan Simester. 1998. On the role of sale signs. *Marketing Sci.* **17**(2) 139–155.

Anderson, Eric, Duncan Simester. 2001. Are sale signs less effective when more products have them? *Marketing Sci.* **20**(2) 121–142.

Anderson, Eric, Duncan Simester. 2004. The cost of inconsistent prices: How do customers react to discounts on items they recently purchased? Working paper, MIT, Cambridge, MA.

Bagwell, Kyle. 1987. Introductory price as a signal of cost in a model of repeat business. *Rev. Econom. Stud.* **54**(3) 365–384.

Bagwell, Kyle, Michael Riordan. 1991. High and declining prices signal product quality. *Amer. Econom. Rev.* **81**(1) 224–239.

Blattberg, Robert, Richard Briesch, Edward Fox. 1995. How promotions work. *Marketing Sci.* **14**(3) G122–G132.

Blattberg, Robert, Scott Neslin. 1990. *Sales Promotion Concepts, Methods, and Strategies*. Prentice-Hall, Englewood Cliffs, NJ.

Dawar, Niraj, Miklos Sarvary. 1997. The signaling impact of low introductory price on perceived quality and trial. *Marketing Lett.* **8**(3) 251–259.

Dekimpe, Marnik, Dominique Hanssens, J. Silva-Risso. 1999. Long-run effects of price promotions in scanner markets. *J. Econometrics* **89** 269–291.

Direct Marketing Association. 1998. *State of the Catalog Industry Report*. DMA, New York.

Direct Marketing Association. 1999. *Statistical Fact Book*, 21st ed. DMA, New York.

Direct Marketing Association. 2000. *State of the Catalog Industry Report*. DMA, New York.

Direct Marketing Association. 2001. *Statistical Fact Book*, 23rd ed. DMA, New York.

Erdem, T., M. P. Keane. 1996. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Sci.* **15**(1) 1–20.

Heckman, James. 1979. Sample selection bias as a specification error. *Econometrica* **47** 153–161.

Inman, Jeffrey J., Leigh McAlister. 1993. A retailer promotion policy model considering promotion signal sensitivity. *Marketing Sci.* **12**(4) 339–356.

Inman, Jeffrey J., Leigh McAlister, Wayne D. Hoyer. 1990. Promotion signal: Proxy for a price cut. *J. Consumer Res.* **17**(1) 74–81.

Jedidi, Kamel, Carl F. Mela, Sunil Gupta. 1999. Managing advertising and promotion for long-run profitability. *Marketing Sci.* **18**(1) 1–22.

Kalyanaram, Gurumurthy, Russell S. Winer. 1995. Empirical generalizations from reference price research. *Marketing Sci.* **14**(3), Part 2 G161–G169.

Krishna, Aradhna. 1992. The normative impact of consumer price expectations for multiple brands on consumer purchase behavior. *Marketing Sci.* **11**(3) 266–286.

Krishna, Aradhna. 1994. The impact of dealing patterns on purchase behavior. *Marketing Sci.* **13**(4) 351–373.

Mela, Carl F., Sunil Gupta, Donald R. Lehmann. 1997. The long-term impact of promotion and advertising on consumer brand choice. *J. Marketing Res.* **34**(2) 248–261.

Mela, Carl F., Kamel Jedidi, Douglas Bowman. 1998. The long-term impact of promotions on consumer stockpiling behavior. *J. Marketing Res.* **35**(May) 250–262.

Milgrom, Paul, John Roberts. 1986. Price and advertising signals of product quality. *J. Political Econom.* **94**(4) 796–821.

Neslin, Scott, Robert Shoemaker. 1989. An alternative explanation for lower repeat rates after promotion purchases. *J. Marketing Res.* **26**(2) 205–213.

Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedict E. M. Steenkamp, Dominique M. Hanssens. 2001. The category demand effects of price promotions. *Marketing Sci.* **20**(1) 1–22.

Pauwels, Koen, Dominique M. Hanssens, S. Siddarth. 2002. The long-term effects of price promotions on category incidence, brand choice, and purchase quantity. *J. Marketing Res.* **39**(4) 421–439.

Rothschild, Michael J., William C. Gaidis. 1981. Behavioral learning theory: Its relevance to marketing and promotions. *J. Marketing* **45**(2) 70–78.

Schmalensee, Richard. 1978. A model of advertising and product quality. *J. Political Econom.* **86**(3) 485–503.

Simester, Duncan. 1995. Signaling price image using advertised prices. *Marketing Sci.* **14**(2) 166–188.

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

Thompson, Patrick A., Thomas Noordewier. 1992. Estimating the effects of consumer incentive programs on domestic automobile sales. *J. Bus. Econom. Statist.* **10**(4) 409–417.

Tirole, Jean. 1988. *The Theory of Industrial Organization*. MIT Press, Cambridge, MA.