PRICE STICKINESS AND CUSTOMER ANTAGONISM[*]

Eric T. Anderson and Duncan I. Simester

Managers often state that they are reluctant to vary prices for fear of "antagonizing customers." However, there is no empirical evidence that antagonizing customers through price adjustments reduces demand or profits. We use a 28-month randomized field experiment involving over 50,000 customers to investigate how customers react if they buy a product and later observe the same retailer selling it for less. We find that customers react by making fewer subsequent purchases from the firm. The effect is largest among the firm's most valuable customers: those whose prior purchases were most recent and at the highest prices.

JEL: E3, D12, E44, C93

Keywords: price stickiness, price adjustments, customer antagonism, field experiments

"It seems essential, therefore, to gain a better understanding of precisely what firms mean when they say that they hesitate to adjust prices for fear of antagonizing customers."

Blinder et al. (1998 page 313)

I.   Introduction

The assumption that nominal prices are sticky is fundamental to Keynesian economics and forms a basic premise of many models of monetary policy. A leading explanation for why prices are slow to adjust is that firms do not want to antagonize their customers. While it has been argued that price adjustments may antagonize customers, there is little empirical evidence that this affects demand or profits. In this paper we study how downward price adjustments can lead to lower demand and lost profits. We show that many customers stop purchasing if they see a firm charge a lower price than they previously paid for the same item. This customer boycott is concentrated among the firm's most valuable customers, which greatly magnifies the cost to the firm. Firms can mitigate these costs by limiting the frequency and/or depth of price adjustments.

The findings are replicated in two separate field experiments conducted in different product categories. The first experiment was conducted with a publishing company. Customers were randomly chosen to receive a test or control version of a catalog (the "Test Catalog"). The versions were identical except that the test version offered substantially lower prices on 36 items (the "Test Items"). The loss of demand and profits was particularly dramatic among customers who had recently paid a higher price for one of those items. Lower prices in the test condition led to 14.8% *fewer* orders over the next 28-months, which equates to over $90 in lost revenue per customer (including revenue from the Test Catalog itself).

Decomposing the results reveals that the price adjustments had two effects. As expected, orders from the Test Catalog were higher in the test condition as some customers took advantage of the discounted prices. However, this short-run effect was overwhelmed by a sharp reduction in orders from other catalogs. Price adjustments in a single catalog had negative spillover effects on future orders (and overall revenue).

To investigate the robustness and generalizability of the results we replicate the key findings with a separate company in a different product category (clothing). This replication is particularly noteworthy as most customers expect to see lower prices on clothing at the end of a season. Despite this, we show that sending a "Sale" catalog to customers in the days immediately after Christmas reduced purchases by customers who had previously paid a higher price for one of the discounted items.

Although the findings confirm that price adjustments can lower demand and profits, we recognize that there may be multiple explanations for this effect. One explanation is that customers are antagonized if they observe the firm charging lower prices than what they previously paid. Another explanation is that lower prices may have prompted customers to update their price expectations and change their purchase behavior. A third possibility is that low prices may have influenced customers' beliefs about product quality.

To evaluate these explanations we exploit heterogeneity in the experimental treatment effects. For example, the reduction in demand and profits was restricted to customers who had recently purchased one of the discounted items and paid a high price in that earlier purchase. Customer

antagonism is consistent with these boundaries. We also find some evidence that customers updated their price expectations and delayed purchasing. However, this does not appear to fully explain the decrease in demand.

Other explanations are ruled out by the randomized allocation of customers to the two experimental conditions. For example, the randomization ensures that competitive reactions, inventory constraints, macroeconomic changes and customer characteristics cannot explain the decrease in demand. These factors were equally likely to affect sales in both conditions.

The results of the field experiments had a direct impact on the pricing policies of the two participating retailers. After learning of the results, the company that provided data for our first experiment responded by no longer sending catalogs containing discounts to customers who had recently purchased one of the discounted items. This effectively reduced the degree to which prices are varied to an individual customer. The company that participated in the second study also responded by restricting price changes. For example, the company removed a discounted item from the back page of a widely circulated catalog to avoid antagonizing approximately 120,000 customers who had previously purchased the item at full price. Even before learning of the findings, this company had policies that limited the frequency of discounts. Managers acknowledged that these policies reflected concern that frequent price adjustments may antagonize customers.

These reactions suggest that firms may be able to mitigate the reduction in demand by charging different prices to different customers. If the products are durables that most customers only purchase once, the cost of foregoing these price changes will be minimal. However, most

retailers are not able to price discriminate as perfectly as the two catalog retailers in this study. When firms cannot charge different prices to individual customers the effects that we report will tend to reduce the optimal frequency and/or depth of price adjustments.

Previous Research

The research on price stickiness can be broadly categorized into two topics: (1) are prices sticky; and (2) why are they sticky? The evidence that prices respond slowly to business cycles is now extensive (Gordon 1990; Weiss 1993). In recent years much of the attention has turned to the second question. One set of explanations argue that costs may be constant, either within a limited neighborhood or within a limited time period. Another common explanation studies the (menu) cost of changing prices. The more costly it is to change prices the less frequently we would expect them to change. Other explanations have considered imperfections in the information available to price setters and asymmetries in the demand curve.

In this paper we study the role of customer antagonism. This explanation argues that firms do not want to change prices as doing so may antagonize their customers. Hall and Hitch (1939) were among the first to empirically investigate this issue. They interviewed a sample of managers to learn how prices are set. The managers' responses included statements such as: "Price changes [are] a nuisance to agents, and disliked by the market," and "Frequent changes of price would alienate customers," (pages 35 and 38). More recently, Blinder, Canetti, Lebow and Rudd (1998) also asked managers why they did not vary their prices. This study was conducted on a large scale and involved 200 interviews with senior executives conducted over a two year period. The most common response was that frequent price changes would "antagonize" or "cause difficulties" for customers, leading the authors to conclude that we need

to better understand the role of customer antagonism (pages 85 and 308, see also the introductory quote at the start of this paper). The findings have since been corroborated by similar studies surveying managers in Canada (Amirault, Kwan and Wilkinson 2004) and a broad range of European countries (Hall, Walsh and Yates 1997; Fabiani et al. 2004; Apel, Friberg and Hallsten 2005).[1] While these studies have raised awareness that customer antagonism may contribute to price stickiness, there are limits to what can be learned from survey data. Notably, the data does not allow researchers to measure whether antagonizing customers through price adjustments reduces demand or profits.

The evidence in this paper is related to Rotemberg's theoretical research on firm altruism. Customers in Rotemberg's models only want to transact with firms that are "altruistic." Although customers interpret firms' actions generously, they boycott firms when there is convincing evidence that their expectations regarding altruism are violated. This reaction has been used to help explain why prices are sticky (Rotemberg 2005); investigate how customers react to brand extensions (Rotemberg 2008); and explore how fairness influences firms' pricing decisions (Rotemberg 2004). We will present evidence that the reductions in future purchases do not just reflect changes in customers' expectations about prices or product quality. Instead, the effects appear to influence how customers view less tangible characteristics of the firm and its brand. This evidence may be seen as direct support for Rotemberg's claim that customers' perceptions of whether firms are altruistic contribute to the value of a firm's brand (Rotemberg 2008).

---

[1] Other empirical evidence includes Zbaracki et al. (2004), who report findings from an extensive study of the cost of changing prices at a large industrial firm. Their findings include a series of anecdotes and quotations indicating that managers at this firm were concerned that frequent price changes would damage the firm's reputation with its customers.

While our findings are clearly consistent with Rotemberg's models, this is not the only

explanation for why customers may stop purchasing from a firm when they are antagonized by

price adjustments.   It is possible that firms are simply risk averse and want to minimize the risk

of future antagonism.[2] Other explanations are also suggested by previous research on the role

of reputations and firm brands (Wernerfelt 1998; Tadelis 1999).

Customer Antagonism When Prices Increase

We measure the response to *downward* price adjustments. This is a natural place to start as it

invokes a strong strawman: lower prices lead to higher sales. There is a complementary stream

of research that studies customer antagonism in response to upwards price adjustments. The

origins of this work can be traced to Phelps and Winter (1970) and Okun (1981). Okun

introduces the label "customer markets" to describe markets with repeated transactions. He

argues that if price increases antagonize buyers, then sellers may respond to unobserved price

shocks by absorbing the cost changes rather than increasing their prices.   Nakamura and

Steinsson (2008) extend this intuition in a recent paper. In their model, price rigidity serves as a

partial commitment device, which enables sellers to commit not to exploit customers'

preferences to repeatedly purchase from the same seller. The commitment mechanism is

endogenous and relies on the threat that a deviation would trigger an adverse shift in customer

beliefs about future prices.[3]

As support for their model, Nakamura and Steinsson (2008) cite two experimental papers.

Renner and Tyran (2004) demonstrate that when buyers are uncertain about product quality

---

[2] We thank an anonymous reviewer for this suggestion.
[3] In a related paper, Kleshchelski and Vincent (2009) investigate how customer switching costs create an
incentive for firms to build market share. They demonstrate that this may prompt firms to absorb a
portion of transitory cost increases without increasing prices.

then sellers are less likely to raise prices in response to cost increases if there is an opportunity to develop long-term relationships. This effect is particularly pronounced if buyers cannot observe the cost increases. The explanation offered for these effects is that the sellers are willing to absorb the additional costs rather than risk antagonizing the buyers. These results complement an earlier experimental study (Cason and Friedman 2002) that investigates how variation in customers' search costs affects both their willingness to engage in repeated relationships and the resulting variation in prices. The authors show that higher search costs tend to result in more repeated relationships, increasing sellers' profits and leading to less variation in prices.

Other Relevant Research

Our findings are also related to previous work on intertemporal price discrimination. Research on the timing of retail sales argues that firms can profit by occasionally lowering prices and selling to a pool of low-valuation customers (Conlisk, Gerstner and Sobel 1984; Sobel 1984, 1991). These arguments do not consider the possibility that customers who observe the lower prices may be less likely to make future purchases. Our findings represent a countervailing force that may limit a firm's willingness to price discriminate.

The findings are also relevant to recent work on price obfuscation. This literature recognizes that many firms adopt practices that make it difficult for customers to compare prices (Ellison and Ellison 2009; Ellison 2006). Price obfuscation may allow firms to mitigate the reactions that we document in this paper. If customers cannot easily compare prices then we would not expect the same outcomes. In this respect, our findings may help to explain the use of price obfuscation mechanisms.

Finally, we can also compare our findings with research on reference prices (Kahneman, Knetsch and Thaler 1986). A challenge in the reference price literature is to identify the correct "reference" price. Some customers may use later prices as a reference against which to evaluate the price paid in earlier transactions. Under this interpretation, the findings are easily reconciled with the reference price literature: customers who later see the firm charging a lower price may conclude they experienced a loss from overpaying in the past. We caution that this definition of the reference price is not unique, and alternative definitions may lead to different predictions.

Plan of the Paper

The paper continues in Section 2 with a detailed description of the experimental setting and the design of the first field experiment. We present initial findings in Section 3, which provide both context and motivation for the results that follow. We investigate how the outcome is moderated by the recency of customers' prior purchases in Section 4. In Section 5 we consider the price paid in the prior purchases, together with the persistence of the results across the 28-month measurement period. In Section 6 we investigate three alternative explanations for the results, and then in Section 7 we present a replication in a different product category. The paper concludes with a summary of the findings and limitations in Section 8.

II.   Study Design

The primary field experiment was conducted with a medium-sized publishing retailer that manufactures and sells a range of approximately 450 products targeted at well-educated retail

customers.[4] All of the products carry the retailer's brand name and are sold exclusively through the company's catalogs. At the time of our study the firm also operated an Internet site, but few customers placed their orders online. The products are durables with similar characteristics to books, computer software and music. It is rare for customers to buy multiple units of the same item (few customers buy two copies of *Oliver Twist*) and so incremental sales typically reflect purchases of other items (buying both *Oliver Twist* and *David Copperfield*). The average inter-purchase interval is 48-weeks.

Interviews with managers revealed that the firm engages in inter-temporal price discrimination, charging relatively high regular prices interspersed with frequent shallow discounts and occasional deep discounts. A review of the firm's historical transaction data confirms that there had been wide variation in prices paid for the same item. Approximately one fifth of transactions occur at the regular price, with most of the remaining transactions occurring at either a shallow (20%-30%) or deep (50%-60%) discount. We report a frequency distribution describing the discounts received in the two years before the Test Catalog was mailed in Figure I.

Figure I about here

Design of the Test Catalog

The field test encompassed two time periods separated by the mailing of the Test Catalog. Our data describes individual customer purchases in the 8-years before the Test Catalog was mailed

---

[4] The results of this study for a small subset of customers were previously described in Anderson and Simester (2004). This earlier paper uses data from this study (and two other studies) to investigate a different research question: comparing how prospective and existing customers react to discounts.

(the "pre-test" period), and the 28-months after this date (the "post-test" period). It will be important to remember that our measure of sales during the post-test period includes purchases from the Test Catalog itself. This allows us to rule out the possibility that the findings merely reflect inter-temporal demand substitution.

There were two different versions of the Test Catalog: a "shallow discount" and a "deep discount" version. A total of 55,047 retail customers were mailed the Test Catalog, all of whom had previously purchased at least one item from the company. Approximately two-thirds of the customers (36,815) were randomly assigned to the shallow discount condition, and the remaining customers (18,232) were assigned to the deep discount condition.[5] The decision to assign a larger fraction of customers to the shallow discount was made by the firm and was outside our control, but does not affect our ability to interpret the results.

We confirm that the allocation of customers to the two conditions was random by comparing the historical purchases made by the two samples of customers. In Table I we compare the average *Recency*, *Frequency* and *Monetary Value* (RFM) of customers' purchases during the 8-year pre-test period.[6] If the assignment were truly random, we should not observe any

---

[5] In Harrison and List's (2004) nomenclature, this is a "natural field experiment". They distinguish between "artefactual field experiments", which are the same as conventional lab experiments but with a non-student subject pool; "framed field experiments", which introduce field context; and "natural field experiments", which also occur in a field context and use subjects who do not know that they are participants in an experiment.

[6] "Recency" is measured as the number of days since a customer's last purchase. "Frequency" measures the number of items that customers previously purchased. "Monetary Value" measures the average price (in dollars) of the items ordered by each customer. The inter-purchase interval (48 weeks) is much shorter than the average of the recency measure (646 days). These measures are not directly comparable: the inter-purchase interval describes the time between purchases, while the recency measure includes some customers who will make no additional purchases.

systematic differences in historical sales between the two samples. Reassuringly, none of the differences are significant despite the large sample sizes.

Table I about here

The Test Catalog was a regularly scheduled catalog containing 72 pages and 86 products. The only differences between the two versions were the prices on the 36 "Test Items". These 36 items were discounted in both versions, but the discounts were larger in the Deep Discount condition. In the Shallow Discount condition, the mean discount on the Test Items was 34%. In the Deep Discount condition, the mean discount was 62%. These yielded mean prices of $133.81 and $77.17 on the 36 Test Items in the two conditions (compared to a mean regular price of $203.83). The size of the price discounts were chosen to be large enough to generate an effect, but not so large that they were outside the range of historical discounts (see Figure I). The prices of the other 50 items were all at their regular prices in both versions.

The prices of the Test Items were presented as "Regularly $x Sale $y" in the deep discount version, and as "Regularly $x Sale $z" in the shallow discount version. The regular price ($x) was the same in both conditions, but the sale price was lower in the deep discount condition ($y < $z). The use of identical text ensured that the discounted price was the only difference between the two versions, and also explains why we used shallow discounts rather than the regular price as a control. There is considerable evidence that customers are sensitive to the word "Sale", even when prices are held constant (Anderson and Simester 2001). Charging the regular price as a control would have made it difficult to distinguish the effect of the price change from the "Sale" effect. By using the same wording in both conditions, we avoid this

confound. As we will discuss, using shallow discounts as a control does not affect our ability to

measure how price adjustments affect sales.

Mailing and Pricing Policies after the Test Catalog

The firm agreed to use the same mailing policy for all customers once the experimental

manipulation was over. To confirm compliance we obtained data describing post-test mailings

for a random sample of 16,271 of the 55,047 customers involved in the test (approximately 30%

of the full sample). Customers in the deep discount condition received a mean of 48.47 catalogs

in the post-test period, compared to 48.77 catalogs in the shallow discount condition. The

difference in these means is not close to statistical significance.

Notice also that the paper's key finding is that sales are lower in the Deep Discount condition

(compared to the Shallow Discount condition). As we will discuss, the deep discounts led to an

increase in sales from the Test Catalog itself. If this affected the firm's subsequent mailing

decisions, it would tend to increase the number of catalogs mailed to customers in this

condition. This could not explain the *decrease* in post-test sales.

It is also useful to consider the firm's pricing policy after the Test Catalog was mailed. The

distribution of discounts received in the post-test period is reported in Figure II. The use of

discounts persisted, with a noticeable increase in the frequency of deep discounts. Most

catalogs mailed during the post-test period contained at least some items with deep discounts.

Because customers in the two conditions received the same downstream catalogs, this change

in policy does not affect our ability to compare the behavior of customers in the two

experimental conditions.

Figure II about here

Predictions

In Figure III we present a histogram of the number of Test Items that the 55,047 customers had

purchased in the 8-year pre-test period. Just over 47% of the customers (25,942) had not

purchased any of the Test Items before receiving the Test Catalog. Of the remaining customers,

very few customers had purchased more than one or two of the 36 Test Items, with less than

0.3% purchasing more than 10 and no customers purchasing every item. As a result, the Test

Catalog offered all customers an opportunity to purchase discounted Test Items that they did

not already own.

Figure III about here

The opportunity to purchase discounted items from the Test Catalog provides a "strawman"

prediction: because our measure of post-test demand includes purchases from the Test Catalog

itself, a standard model would predict higher demand in the Deep Discount condition (which

offered lower prices). We label this the "Low Prices" prediction:

Low Prices

Post-test demand is <u>higher</u> in the Deep Discount condition than in the Shallow Discount

Condition.

Customer antagonism suggests an alternative prediction. Customers in the Deep Discount condition were more likely to see prices in the Test Catalog that were lower than what they paid for the same item. If this antagonizes customers then we might expect lower sales in this condition. We label this the "antagonism" prediction:

Antagonism

Post-test sales are <u>lower</u> in the Deep Discount condition than in the Shallow Discount Condition.

This prediction requires only a relatively simple model of customer behavior. It is sufficient that if customers see lower prices than what they paid then they are less likely to make additional purchases. Because this is more likely to occur in the Deep Discount condition we expect fewer sales in that condition.[7]

We will also consider three interactions. The first interaction distinguishes the 29,105 customers who had purchased one of the 36 Test Item before receiving the Test Catalog from the 25,942 customers who had not. We only expect customers to be antagonized if they see lower prices on items that they had purchased. Therefore, we expect a more negative (less positive) reaction to the deep discounts among the 29,105 customers with a prior purchase:[8]

---

[7] Notice that we do not need to rule out favorable responses if customers see a higher price than what they previously paid. We expect lower post-test sales in the Deep Discount condition irrespective of whether there is a positive response (or no response) to seeing higher prices. However, in Section 4 we do investigate how the price that customers previously paid affects their response to the deep discounts.

[8] It is possible that customers may be concerned about whether other customers were antagonized. However, previous studies have consistently reported that decision-makers show little concern for the sufficiency of other customers' outcomes (Guth and van Damme 1998; Selten and Ockenfels 1998).

Past Purchase Interaction

The deep discounts have a more negative impact on post-test demand among customers who had previously purchased one of the Test Items.

There is an alternative explanation for this interaction. Because customers are unlikely to purchase the same item twice, customers who had already purchased may have been less likely to take advantage of the deep discounts in the Test Catalog. We will investigate this explanation in Section 6, together with other alternative explanations.

Customer antagonism may also depend upon how recently customers had purchased. Customers who purchased a long time ago may find it harder to remember the price that they paid. Moreover, those who can remember the price may experience less regret about not waiting because they have had additional opportunities to consume. We will investigate whether the time between a customer's prior purchase and the mailing date for the Test Catalog contributes to the outcome:

Time Interaction

The deep discounts have a less negative effect on post-test sales if the time since a customer's previous purchase is longer.

We also investigate whether there is evidence that customers who paid higher prices for Test Items were more antagonized by the deep discounts. This "past price" interaction is somewhat complicated by the experimental design and so we will delay further discussion of it until we present the results.

III.   Initial Results

In this section we present initial findings that provide both context and motivation for the main analysis that follows. To begin we ask how the deep discounts affected post-test purchases by the full sample of 55,047 customers. These results are reported in Table II, where we describe the average number of orders placed by customers in the Deep and Shallow Discount conditions. We distinguish between orders from the Test Catalog itself, and orders from other catalogs (not the Test Catalog) during the post-test period.[9] We also present the average revenue earned, the average size of their orders and the average price of the items that customers purchased.

Table II about here

Because this overall comparison aggregates customers for whom the outcome was positive with others for whom it was negative, it reveals few significant differences between the two samples. We can illustrate this by returning to the motivating example used in the Introduction. In Table III we focus on customers who paid a high price (above the shallow discount price) for a Test Item in the 3-months before the Test Catalog was mailed. In the Introduction we anticipated that these were the customers most likely to have been antagonized by the deep discounts.

---

[9] When customers call to place an order, they are asked for a code printed on the back of the catalog. This code is also printed directly on the mail-in order form.   The data we received contain the catalog code for each order, so we can identify orders from the Test Catalog and orders from other catalogs. There are a small number of orders for which the catalog code is not available, including some Internet orders. Fortunately these instances are rare (there were very few Internet orders during the period of the study), and have little effect on the findings.

Table III about here

The deep discounts resulted in fewer post-test orders from these customers. In the Deep

Discount condition customers placed an average of 2.78 post-test orders compared to 3.27

orders in the Shallow Discount condition. The difference (0.48 orders or 14.8%) is statistically

significant and can be fully attributed to the prices in the Test Catalog. The findings also help

clarify the source of the effect. The differences result solely from changes in the number of

orders placed, rather than changes in the composition of those orders; there is no difference

between conditions in either the average number of items per order or the average prices of the

items purchased. The reduction in orders at least partly reflects an increase in the proportion of

customer who placed no orders during the post-test period (34.0% in the Deep Discount

condition versus 27.1% in the Shallow Discount condition). It is this result that led us to describe

the effect as a customer "boycott".

The decrease in orders is a strong result: merely sending these customers a catalog containing

lower prices reduced purchases. The effect is large, and is precisely the outcome that managers

appear to anticipate when stating they were reluctant to adjust prices for fear of antagonizing

their customers (Blinder et al. 1998; Hall and Hitch 1939). Notice that customers who had

purchased recently (Table III) are systematically more valuable than the average customer in the

study (Table II). Although they place similar sized orders, the recent purchasers order a lot more

frequently in the post-test period. Because they are more valuable, the cost to the firm is

greatly amplified. Although confidentiality restrictions prevent us from reporting detailed profit

results, sending the deep discount version to these 1,392 customers would have lowered the

profits earned from these customers by over $93,000 (compared to sending the shallow

discount version). To put this in context, the printing and mailing cost for sending the Test

Catalog is approximately 50-cents, or $27,500 to send it to all 55,047 customers in the study.

It is also helpful to remember factors that cannot explain this result. Most importantly, the

difference in orders cannot be explained by differences in the customers themselves. Our

analysis compares customers in the two experimental conditions and random assignment

(which we verified in Table I) ensures that there are no systematic differences between

customers in the two conditions. We can also see that the difference in post-test orders is not

due to a difference in the number of orders from the Test Catalog itself. Orders from the Test

Catalog represent only a very small proportion (less than 3%) of overall post-test demand.

Moreover, orders from the Test Catalog were actually slightly higher in the Deep Discount

condition, presumably due to the lower prices. Finally, because our measure of sales includes

purchases from the Test Catalog, the result cannot be explained by intertemporal demand

substitution (forward-buying). Acceleration in purchases to take advantage of the Test Catalog

discounts would not affect this measure of total orders.

One explanation for the results, which does not depend on customer antagonism, is that the

deep discounts may have changed customers' expectations about the availability of future

discounts. If customers delayed their future purchases in anticipation of these discounts it could

explain a reduction in post-test sales. As we acknowledged in the Introduction, we cannot

completely rule out this explanation. However, in Section 6 we will present results that suggest

this cannot be a complete explanation for all of the findings. In particular, if customers were

waiting for future discounts, the decrease in post-test sales should be larger at prices that do

not offer large discounts. It is these purchases that we would expect customers to forgo while

waiting for additional discounts. However, we find that sales decrease even when the discounts are large.

The initial analysis in this section focused on the sample of customers who we expected to be most susceptible to the antagonism prediction. In the next section we extend the focus to all of the customers in the study, and measure how the effect was moderated by whether customers had previously purchased a Test Item and (if so) how recently they had purchased it.

IV.   The Past Purchase and Time Interactions

To directly estimate the *Past Purchase* and *Time* interactions we use a multivariate approach. Because our initial analysis revealed that the primary effect is upon the number of orders placed, rather than the composition of those orders, we use a "count" of orders placed as the dependent variable (we later also use revenue and profits as dependent variables). We estimate this variable using Poisson regression, which is well-suited to count data.[10] Under this model the number of orders from customer i (Q_i) is drawn from a Poisson distribution with parameter λ_i:

(1) $\text{Prob}(Q_i = q) = \dfrac{e^{-\lambda_i} \lambda_i^q}{q!},$ where q=0, 1, 2, ..., and $\ln(\lambda_i) = \boldsymbol{\beta} \mathbf{X}_i$.

To estimate how the outcome was moderated by the time interaction since a customer's prior purchase, we use the following specification:

(2) $\boldsymbol{\beta} \mathbf{X}_i = \beta_1 Deep\,Discount_i + \beta_2 Deep\,Discount_i * Time_i + \beta_3 Time_i + \theta Z_i$

These variables are defined as:

---

[10] The results are also robust to using Zero Inflated Poisson (ZIP) and OLS specifications.

Deep Discount$_i$ A binary variable indicating whether customer *i* was in the deep discount condition.

Time$_i$ The log of the number of months between the Test Catalog mailing date and customer *i*'s most recent purchase of a Test Item.

For completeness the vector *Z* includes the log of the historical *Recency*, *Frequency* and *Monetary Value* (RFM) measures as control variables. Because these variables do not vary systematically across the two conditions (Table I), their inclusion or absence has little impact on the coefficients of interest (β_1 and β_2). Among customers who previously purchased a Test Item, β_1 describes how receiving the deep discounts affected post-test orders by a "benchmark" customer, who purchased a Test Item immediately before receiving the Test Catalog (*Time* equals zero). As the time since the customer's prior purchase increases, the estimated impact of the deep discounts is moderated by β_2. The time interaction predicts that β_2 will have a positive sign: the longer the time since the prior purchase, the smaller the reduction in post-test sales.

Notice that this model preserves the benefits of the randomized experimental design. The coefficients of interest measure the percentage difference in post-test sales between customers who received the deep and shallow discount versions of the Test Catalog. As a result, the coefficients of interest cannot be explained by difference in customer characteristics between the two experimental treatments, or by intervening competitive or macroeconomic events. We rely heavily on this feature of the study as it allows us to rule out a wide range of alternative explanations. The model is estimated separately on the 28,642 customers who had previously purchased a test item and on the 26,405 customers who had not previously purchased a test

item.[11] Coefficients and standard errors for both models are reported in Table IV.

Table IV about here

Column 1 focuses on customers with a prior Test Item purchase. For these customers β_1 is

negative and significant, indicating that for our benchmark customer (for whom *Time* equals

zero) the deep discounts led to a 7.8% reduction in post-test sales. This replicates our univariate

results and confirms that merely sending customers a catalog containing lower prices led to

fewer orders by these customers. For the same customers β_2 is positive and significant. This

result is consistent with the *Time* interaction, and confirms that the loss of sales is smaller when

more time has passed since a customer's earlier purchase.

The findings also reveal that the drop in post-test sales was limited to customers who had

previously purchased one of the discounted items. For customers who had not previously

purchased a Test Item (Column 2) the difference in sales between the two conditions was not

significant.   This is consistent with the *Past Purchase* prediction, which anticipated that the

deep discounts would only reduce sales among customers who had previously purchased one of

the Test Items.[12]

---

[11] Joint estimation of the two models yields the same pattern of results but the findings are more difficult
to interpret (they require three-way interactions). For customers who had not previously purchased a Test
Item we calculate the *Time* since their most recent purchase of any item. This is the same as our *Recency*
measure and so we omit *Recency* from this model.

[12] We did not expect these customers to be antagonized by the deep discounts, but we would have
expected them to take advantage of the lower prices. Further investigation confirms that when we
restrict attention to orders from the Test Catalog itself, we see a strong *increase* in sales in the Deep
Discount condition, but this was offset by lower sales from subsequent catalogs.

We can use the coefficients in Table IV to calculate how much time is required to elapse for the deep discounts to have a positive effect on post-test sales. Given coefficients of -0.078 for β_1 and 0.026 for β_2, the net impact of the deep discounts equals zero when *Time* is equal to 20.1 months.[13] We conclude that the estimated impact of the deep discounts was negative if customers purchased a Test Item within approximately two years of receiving the Test Catalog.

This time interaction has an important additional implication for the firm. Recall that the initial findings in the previous section (and the *Recency* and *Time* coefficients in Table IV) confirm that customers who purchased recently are systematically more valuable - they purchase significantly more frequently in the post-test period. Because the reduction in post-test sales was focused on these recent customers, there was a disproportionate effect on overall sales. Specifically, 13,830 customers purchased a Test Item within 20.1 months of receiving the Test Catalog. These customers represent just 25% of the sample but contributed approximately 52% of post-test revenue. If the firm mailed the deep discount versions (rather than shallow discount version) to each of these 13,830 customers its profits would decrease by approximately $155,000.[14]

Equation (2) imposes a functional form on the interaction between the *Time* and *Deep Discount* variables. We can relax this restriction by grouping customers into segments based on the time since their earlier Test Item purchase, and directly estimating the impact of the deep discount on each segment. In particular, we group customers into four segments based on the timing of their earlier purchases: less than 250 days; 250 to 500 days; 500 to 750 days; or over 750 days since the Test Catalog was mailed. We then estimate the following Poisson regression model for

---

[13] Calculated as $e^{\frac{0.078}{0.026}}$ (recall that *Time* is measured in months and has a log specification).

[14] Calculated as the difference in the average profit in each condition, multiplied by the 13,830 customers.

each segment:

$$(3) \qquad \beta \mathbf{X_i} = \beta_1 Deep\, Discount_i + \theta Z_i$$

The β_1 coefficients for each segment are summarized in Figure IV and complete findings are provided in Table V.[15] The deep discounts led to a significant loss of sales (-5.22%) from the 4,413 customers who had purchased within 250 days of receiving the Test Catalog, and a significant increase in sales (+7.87%) from customers whose prior purchases occurred over 750 days ago. The results for the other two segments fall between these two results.   This consistent monotonic relationship is reassuring, and confirms that we cannot attribute the role of time solely to the functional form imposed in Equation (2).

Table V and Figure IV about here

The positive outcome for customers whose prior purchases were not recent (over 750 days ago) is worthy of comment.   We offer two possible explanations. First, recall that customers had generally purchased at most one or two of the Test Items that were discounted in the Test Catalog (see Figure III). The deep discounts in the Test Catalog gave them the opportunity to purchase other items at low prices. This motivated our *Low Prices* prediction that sales would be higher in the Deep Discount condition. Second, it is possible the deep discounts persuaded some customers that the firm offers low prices, which may have prompted them to pay more attention to future catalogs (Anderson and Simester 2004). We investigate both explanations in

---

[15] In the next set of analysis we show that the effect is concentrated among customers who paid a high price in their previous purchase. We restrict attention to these customers in this analysis.

Table VI, where we distinguish orders from the Test Catalog itself and post-test orders from

other catalogs.

Table VI about here

The deep discounts led to more orders from the Test Catalog for all customer segments. This is

consistent with customers taking advantage of the deep discounts in that catalog to purchase

items they had not already purchased (the *Low Prices* prediction). The effect is particularly

strong for customers whose prior purchases were less recent. These are the same customers for

whom we observe a positive outcome in Figure IV. Among customers who had not purchased

for over 750 days, the positive effect of the deep discounts extends beyond the Test Catalog to

also increase orders from other catalogs by 6.6%. This is consistent with our second

explanation: more favorable price expectations may have expanded the range of occasions on

which these customers searched for products in the firm's catalogs.

Summary

We conclude that the findings offer strong support for the *Past Purchase* and *Time* interactions.

The reduction in sales is limited to customers who had previously purchased a Test Item and is

stronger for customers who purchased more recently. We next consider how the price that

customers paid in their earlier purchases influences the outcome. We will then investigate

whether the findings persist throughout the post-test period, and whether they survive when

we use post-test measures of revenue and profit as the dependent variable.

V.   Additional Results

If price variation leads to customer antagonism, we might expect that customers who paid the highest prices in their earlier purchases would be most antagonized. There is considerable variation in the prices that customers paid for Test Items before receiving the Test Catalog. We can illustrate this variation by using this past price to group customers into three segments, which we label "Low", "Medium" and "High":

Segment	Past Price Level	Number of Customers
Low	Less than the deep discount price	463
Medium	Between the deep and shallow discount prices	12,633
High	Above the shallow discount price	16,009

We do not expect the 463 customers who paid less than the deep discount price to be antagonized by the deep discounts. For the 12,633 customers in the Medium segment who paid between the deep and shallow discount prices, the possibility of antagonism only arises in the Deep Discount condition, as this was the only condition where prices were lower than what customers previously paid. In contrast, all of the 16,009 customers in the High segment paid above the shallow discount price and saw lower prices when they received the Test Catalog. These customers may be antagonized in both experimental conditions.

Without knowing how customers respond to small and large price differences, we cannot make a clear ex-ante prediction about how the outcome will vary across the High and Medium segments. We can illustrate this uncertainty by considering two extreme examples. If

customers have the same negative reaction to any observed price reduction, then customers in

the High segment will have the same reaction in both experimental conditions. In contrast,

customers in the Medium segment will react negatively in the Deep Discount condition but not

in the Shallow Discount condition. Our comparison of the two conditions will reveal a negative

response in the Medium segment and no effect in the High segment. A second (equally

extreme) behavior is that customers only react to large price differences. It is possible that none

of the customers in the Medium segment observe a large enough price difference to prompt a

reaction, while the outcome in the High segment depends upon the amount customers paid.

Customers who paid the most may react in both conditions, while others may only react in the

Deep Discount condition. Comparison of the two conditions will reveal no effect in the Medium

segment and a negative effect for some customers in the High segment.

We can investigate this issue empirically. Our experimental design ensures that within each

segment customers were randomly assigned to the Deep and Shallow Discount conditions.

Therefore, we can measure how the deep discounts affected post-test sales in each segment by

estimating Equation (3) separately for each segment. The coefficients, which are reported in

Table VII, reveal clear evidence that the past price plays an important role. Customers in the

High segment reacted adversely to the deep discounts, but we do not observe any reaction in

the Medium segment.   Together, these results suggest that while customers react negatively to

large price differences, they may be willing to overlook small price differences.

Table VII about here

These results also have an important implication for the firm. Customers who pay higher prices tend to be systematically more valuable: the 16,009 customers in the High segment represent approximately 29% of the sample but contribute 47% of the post-test profit. The concentration of the effect amongst these customers amplifies the importance of the effect on the firm's profits.

We conclude that the reduction in demand is concentrated in the segment of customers who had previously paid a high price (above the shallow discount price) for a Test Item. We will focus on this segment in the remainder of our analysis.   We next consider whether the findings persist throughout the post-test period, and then investigate whether they survive when we use post-test measures of revenue and profit as the dependent variable.

Does the Adverse Outcome Persist?

To evaluate the persistence of the results we divided the 28-month post-test period into two fourteen-month sub-periods. We estimate Equation (2) on both sub-periods and report the findings in Table VIII (Columns 1 and 2). Comparing β_1 and β_2 between the two data periods reveals no significant difference in either coefficient (they are also not significantly different from the findings reported for these customers in Table VII). However, the results do imply a different time cutoff for which there is a negative response to the deep discounts. At the start of the post-test period the negative response extends to customers who purchased within 23.5 months of receiving the Test Catalog. At the end of the period the results indicate that the response is only negative for customers whose prior purchase occurred with 9.7 months of the Test Catalog. We conclude that although the negative effect survives for more than a year after the Test Catalog was mailed, there is evidence that the effect decays over time.

Profit and Revenue

The regression results reported so far consider only the number of orders purchased during the post-test period. To evaluate robustness we also analyzed two additional dependent measures: the *Total Revenue* and *Total Profit* earned from each customer during the post-test period (where profit is calculated as revenue minus cost of goods sold). Both variables are continuous rather than count measures and so we estimate the models using OLS. The findings are reported in Table VIII (Columns 3 and 4). They confirm that deep discounts lead to a reduction in both revenue and profits. This effect diminishes as the time since a customer's previous purchase increases. The interval at which the net effect is zero is approximately 22 months for both metrics.

Table VIII about here

Summary

We conclude that lower prices leads to fewer purchases by some customers. This effect is strongest amongst customers who had recently paid a high price to buy an item on which the price is later lowered. Unfortunately these include many of the firm's most valuable customers, magnifying the importance of the effect. Readers who are solely interested in the existence of this effect may want to read ahead to Section 7, where we replicate the findings with a different company in a different product category. However, other readers may be interested in reviewing alternative explanations for these outcomes. We address this issue in the next section.

VI. Alternative Explanations

The findings reported in the previous sections are consistent with customer antagonism. In this section we evaluate three alternative explanations for the findings. We begin by considering the possibility that customers delayed their purchases in anticipation of future discounts. We then consider the role of both quality signals and demand depletion.

Delayed Purchases

Coase recognized that customers may respond to inter-temporal discounts by delaying their purchases in anticipation of future discounts (Coase 1972). In our context, the discounts in the Deep Discount condition may have alerted customers to the possibility of future discounts. If customers responded by delaying their subsequent purchases, this might explain the reduction in post-test orders. As we have already acknowledged, we cannot fully rule out this explanation, but we can investigate it using several approaches.

First, if customers were waiting for future discounts, the decrease in post-test sales should be larger at prices that represent either small discounts or no discounts. It is these purchases that we would expect customers to forgo while waiting for larger discounts. To investigate this prediction we re-calculate post-test sales using different discount thresholds. For example, "Discounts of at least 60%" counts the number of units in the post-test period at discounts of at least 60%.[16] We re-estimated Equation (2) using eight different discount thresholds and report

---

[16] Notice that we shift from "orders placed" to "units purchased" because discounts are defined at the unit level.

the results from each of these eight models in Table IX.

<center>Table IX about here</center>

Though not statistically significant, there is weak evidence of a trend in the results. At higher

discount thresholds the *Deep Discount* coefficients are slightly less negative, which is consistent

with customers waiting for larger discounts. Moreover, the implied *Time* at which the effect

switches from negative to positive becomes shorter (indicating that we only observe a negative

effect among customers who purchased immediately before receiving the Test Catalog).

However, the findings also confirm that sales decrease even for deeply discounted items, which

is difficult to reconcile with customers waiting for future discounts.

Our second approach recognizes that deeply discounted prices were not unique to the Test

Catalog. Over 38% of purchases in the pre-test period were made at discount levels of at least

50% (see Figure I).[17] If customers had purchased at deep discounts in the past, it seems less

likely they would be surprised by the deep discounts in the Test Catalog, or that they would

change their behavior by delaying future purchases. Therefore, we can evaluate this alternative

explanation by investigating whether the decrease in demand was smaller among customers

who had purchased at deep discounts in the pre-test period. We do so by including a measure

of past discounts in our model:

$$(4) \qquad \beta\mathbf{X}_i = \beta_1 Deep\,Discount_i + \beta_2 Deep\,Discount_i * Time_i + \beta_3 Time_i$$
$$+ \beta_4 Deep\,Discount_i * Maximum\,Past\,Discount_i + \beta_5 Maximum\,Past\,Discount_i + \theta\mathbf{Z}_i$$

---

[17] Many pre-test purchases were made at even deeper discounts (14.6% were made at a discount of at least 60%). Recall that we used a 60% discount level in the Deep Discount condition because it was consistent with past discount levels.

where *Maximum Past* Discount$_i$ measures the largest percentage discount <u>on any item</u> that

customer *i* purchased during the pre-test period. We report the coefficients for this model in

Table X. We also include an alternative model in which *Maximum Past Discount$_i$* is replaced with

a binary variable (*60% Past Discount$_i$*) identifying customers who had previously purchased at a

discount of at least 60%.

Table X about here

Neither of these interaction terms approaches statistical significance. The response to the deep

discounts was apparently not affected by the size of the discounts that customers had received

in the past. This is not what we would expect if the Test Catalog changed customers' price

expectations; the effect should be strongest for customers who had not previously purchased at

a large discount.

We make a final observation about this explanation. At the start of Section 3 we reported the

average price of the items purchased in the post-test period (Table III). Increased willingness to

wait for discounts should lead to customers paying lower average prices. However, the average

price paid (per unit) was not significantly different between the two conditions.

We conclude that delay in anticipation of future discounts does not appear to be a complete

explanation for the results.

Quality Signals

It is possible that customers interpreted the deep discounts on the Test Items as a signal that these items were of inferior quality. To investigate this possibility, we can restrict attention to non-Test Items ("other items"). We do not expect discounts on the Test Items to signal information about the quality of these other items, and so if the effect persists for these other items, it is unlikely to be due to an adverse quality signal.

The Test Items accounted for only 22% of the 94,487 post-test purchases, so the remaining 78% of the purchases were for the approximately 400 other items sold by the firm. We repeat our earlier analysis when distinguishing between post-test sales for Test Items and other items and report the results in Table XI. The pattern of findings is unchanged. Indeed, the findings are stronger for these other items than for the Test Items, presumably because some customers took advantage of the deep discounts (on the Test Items) in the Test Catalog. We conclude that the decrease in sales cannot be fully explained by customers using the deep discounts as a signal that the Test Items are poor quality.

Table XI about here

An alternative interpretation is that the deep discounts lowered the perceived quality of all of the products sold by the firm. This is a relatively implausible explanation in this setting. Most customers in the study had made multiple previous purchases and had received a large number of catalogs from the firm. On average, each customer had purchased approximately 3.3 units prior to receiving the Test Catalog. This increases to 5.2 units among customers who had purchased a Test Item within 20.1 months of receiving the Test Catalog. Moreover, customers'

first purchases on average occurred over two years before they received the Test Catalog and in the intervening period they had received many of the firm's catalogs (approximately once every two to four weeks). Given this extensive experience with the firm, it seems unlikely that customers' overall perceptions of the quality of the firm's products would be affected by the prices in a single catalog.

We conclude that it is unlikely that the results can be explained by changes in customers' expectations about product quality. It also appears unlikely that customers were merely waiting in anticipation of future discounts. It is for this reason that we suggested in the Introduction that the reputation effects reflect intangible brand attributions, rather than inferences about future prices or product quality. We next consider an alternative explanation for the *Past Purchase* and *Time* interactions.

Demand Depletion

Because customers are unlikely to buy the same item twice, those who had already purchased a Test Item had fewer opportunities to take advantage of the deep discounts on these items in the Test Catalog. Similarly, customers with *recent* purchases may have had their immediate needs satisfied, which may also have diminished demand for discounted items in the Test Catalog.

It is helpful to consider the limits of this explanation. Demand depletion may explain why there was *no increase* in sales in the deep discount condition (compared to the shallow discount condition): customers could not take advantage of the lower prices because their demand was depleted. However, the two versions were mailed to equivalent groups of customers (whose demand was equivalently depleted) and so demand depletion cannot explain why sales

decreased in the deep discount condition. Notice also that orders from the Test Catalog account for less than 3% of the post-test orders.   If we restrict attention to orders from catalogs other than the Test Catalog, the findings survive and are actually strengthened by omission of the Test Catalog orders. This result cannot be explained by a diminished response to the Test Catalog itself.[18] We conclude that depletion of demand for Test Items from the Test Catalog cannot fully explain why the deep discounts reduced aggregate post-test orders from these customers.

In the next section we present the findings from a second study conducted with a clothing retailer. These findings offer an opportunity to replicate the results in a different company and product category.

VII. Replication

To investigate whether the effects generalize we conducted a second study with a separate catalog firm that sells private label clothing. The study investigates the impact of mailing a "Sale" catalog containing discounts in the days immediately after Christmas, a period when many customers expect clothing to be discounted. If we can show that the key findings extend to this setting we can be more confident that customer antagonism plays an important role in contributing to price stickiness.

---

[18] Similarly, diminished demand from the Test Catalog cannot explain the findings in the second half of the post-test period. Over 99.7% of purchases from the Test Catalog occur within the first fourteen months of the post-test period.

Approximately 110,000 customers who had previously purchased from the firm were randomly

assigned to equal-sized Test and Control conditions.[19] Those in the Control condition received a

16-page "Sale" catalog containing 132 items, all of which were discounted below their regular

prices. The average regular price of the 132 items was $67.24 and the average discount was

45%. Customers in the Test condition were not mailed this catalog. We measured (post-test)

sales over the subsequent 30-months, including purchases from the Sale catalog.

Before describing these findings, we recognize that there are two important differences in the

experimental manipulation compared to the previous study. First, it was standard for this

company to mail a catalog containing discounts immediately after Christmas, and so the

experimental manipulation was *suppression* of the Sale catalog. Because lowering the price was

standard behavior, it appears implausible that merely receiving this Sale catalog changed

customers' expectations about the frequency of future discounts. The second difference is that

in the previous study all customers received the Test Catalog; we merely varied prices between

the two conditions. This ensured that any effects could be attributed to variation in prices and

not to other information in the catalog. In this second study, the variation is sending the Sale

catalog itself. As a result, we cannot distinguish which feature(s) of the Sale catalog prompted

the observed outcomes.[20]

Table XII about here

---

[19] Comparison of historical purchases confirmed that the assignment of customers to these two
conditions was random.

[20] If customers in the Test condition (who did not receive the Sale catalog) called to order one of the items
in the Sale catalog, they were also charged the discounted price. Therefore, under a strict interpretation
the experiment manipulates the information that customers received about price changes, rather than
the prices themselves. In a catalog setting this distinction is less important because customers cannot
respond to price changes that they do not know about.

The findings are reported in Table XII. We report three models using different samples of customers. In Column 1 we estimate the model using customers who had previously purchased one of the items in the Sale catalog and paid more than the Sale catalog price.   In Column 2 we use customers who had previously paid a price equal to or less than the price in the Sale catalog, while in Column 3 we use customers who had not previously purchased one of the items in the Sale catalog.

Customers who had previously paid a higher price for one of the discounted items (Column 1) made 2.4% *fewer orders* if they received the Sale catalog. This is again a strong result. Our measures include purchases from the Sale catalog, and the drop in sales can be fully attributed to merely receiving the Sale catalog containing lower prices. Customers who had <u>not</u> previously purchased one of the discounted items (Column 3) placed 1.6% *more orders* during the 30-month post-test period if they received the Sale catalog. These outcomes are both significantly different from zero and significantly different from each other (p<0.01). They suggest that customer antagonism contributed to a net reduction in orders of 4.0% (the difference between a 1.6% increase and a 2.4% decrease).

Summary

The results in this replication may initially appear counter-intuitive. Customers would not be surprised to receive a Sale catalog in the days after Christmas, and so we might not expect them to react adversely to the low prices in this catalog. On the other hand, customers do not know *which* products will be discounted in the Sale catalog, and it is likely that they experienced regret if they saw that they could have purchased the same item at a lower price. This is

precisely the situation anticipated by Rotemberg (2004). Recall that customers in his model only want to transact with firms that are "altruistic." An altruistic firm would not allow customers to experience regret, and so regret represents evidence that a firm is not altruistic.

The findings in this second study offer reassurance that the effects are replicable. They also provide evidence that they generalize to other markets, including markets in which customers recognize that prices vary over time. Without this evidence it would be more difficult to claim that the effects contribute to price stickiness to an extent that is relevant to our understanding of monetary policy.

VIII.   Discussion and Conclusions

While customer antagonism is recognized as a possible explanation for price stickiness, there has been virtually no empirical evidence that price adjustments can cause customer antagonism and reduce demand. We present findings from two large-scale field tests that investigate how customers respond to downwards price adjustments. The findings reveal that many customers stop purchasing if a firm charges a lower price than what they previously paid for the same item. We characterize the loss in demand as a customer "boycott" of the firm. The loss in profits was sufficient to cause the firms that participated in our two field experiments to reduce the frequency of price adjustments.

Both of the studies reported in the paper focus on products that are durables, for which repeat purchases are rare. It is possible that the results may be different if we had used products that consumers purchase repeatedly. Customers may be less antagonized by lower prices on items

they have already purchased if they can take advantage of the discounts. However, it is important to remember that the focus on durables in these two studies did not prevent customers from taking advantage of the discounts. Most customers had previously purchased very few of the discounted items, and so they could take advantage of the discounts by purchasing other items.   Indeed, many customers did take advantage of the discounts: in the first study purchases from the Test Catalog itself were over 60% higher in the Deep Discount condition, while in the second study many customers who received the Sale catalog purchased from it.

Our measures may underestimate how customers react to seeing lower prices on items that they have previously purchased.   When customers receive catalogs many of them simply throw them out without opening them. Random assignment ensures that for the Test Catalog this is likely to have occurred at equivalent rates in both conditions. However, if customers did not open the Test Catalog then the variation of prices within the catalog could not have affected their behavior.   This will tend to weaken the aggregate differences between the two treatments and under-estimate the individual effects on the customers who did see the discounts.

While the data and the randomized experimental design allowed us to rule out many alternative explanations, we recognize that there are limitations. The most important limitation is that we do not have access to direct measures of customers' psychological reactions. This absence of intermediate process measures is a limitation common to many field studies.

Northwestern University

MIT

References

Amirault, D., C. Kwan, and G. Wilkinson, "A Survey of the Price-Setting Behavior of Canadian Companies," *Bank of Canada Review*, 2004, 29-40.

Anderson, Eric T., and Duncan I. Simester, "Are Sale Signs Less Effective When More Products Have Them?" *Marketing Science*, XX (2001), 121-142.

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

Apel, M., R. Friberg, and K. Hallsten, "Micro Foundations of Macroeconomic Price Adjustment: Survey Evidence from Swedish Firms," *Journal of Money, Credit and Banking*, XXXVII (2005), 313-338.

Blinder, Alan S., Elie R. D. Canetti, David E. Lebow, and Jeremy B. Rudd , *Asking About Prices: A New Approach to Understanding Price Stickiness* (New York, NY: Russell Sage Foundation, 1998).

Cason, T. N., and D. Friedman, "A Laboratory Study of Customer Markets," *Advances in Economic Analysis and Policy*, II (2002), Article 1.

Coase, Ronald H., "Durability and Monopoly," *Journal of Law and Economics*, XV (2001), 143-149.

Conlisk, John, Eitan Gerstner, and Joel Sobel, "Cyclic Pricing by a Durable Goods Monopolist," *The Quarterly Journal of Economics*, XCIX (1984), 489-505.

Ellison, Glenn, "Bounded Rationality in Industrial Organization," in *Advances in Economics and Econometrics: Theory and Applications*, R. Blundell, W. Newey, and T. Persson, ed., Ninth World Congress (Cambridge UK: Cambridge University Press, 2006).

Ellison, Glenn, and Sara Fisher Ellison, "Search, Obfuscation, and Price Elasticities on the Internet," *Econometrica*, LXXVII (2009), 427-452.

Fabiani, S., M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Matha, R. Sabbatini, H. Stahl, and A. Stokman, "The Pricing Behavior of Firms in the Euro Area: New Survey Evidence," Paper Presented at Conference on "Inflation Persistence in the Euro Area" at the European Central Bank, 2004.

Gordon, Robert J., "What is New-Keynesian Economics?" *Journal of Economic Literature*, XXVIII (1990), 1115-71.

Guth, Wener, and Eric van Damme, "Information, Strategic Behavior and Fairness in Ultimatum Bargaining: An Experimental Study," *Journal of Mathematical Psychology*, XLII (1998), 227-47.

Hall R. L., and C. J. Hitch, "Price Theory and Business Behavior," *Oxford Economic Papers*, II (1939), 12-45.

Hall, S., M. Walsh, and T. Yates, "How do UK Companies Set Prices?," *Bank of England Quarterly Bulletin*, 1997, 180-192.

Harrison, Glenn W., and John A. List, "Field Experiments," *Journal of Economic Literature*, XLII (2004), 1009-1055.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, LXXVI (1986), 728-741.

Kleshchelski, Isaac, and Nicolas Vincent, "Market Share and Price Rigidity," *Journal of Monetary Economics*, LVI (2009), 344-352.

Nakamura, Emi, and Jón Steinsson, "Price Setting in Forward-Looking Customer Markets," Columbia University Working Paper, 1998.

Okun, A. M., *Prices and Quantities: A Macroeconomic Analysis* (Washington DC: Brookings Institution, 1981).

Phelps, E. S., and S. G. Winter, "Optimal Price Policy Under Atomistic Competition," in *Microeconomic Foundations of Employment and Inflation Theory*, E. S. Phelps, ed. (New York, NY: W. W. Norton, 1970).

Renner, Elke, and Jean-Robert Tyran, "Price Rigidity in Customer Markets: An Empirical Study," *Journal of Economic Behavior and Organization*, LV (2004), 575-593.

Rotemberg, Julio J., "Fair Pricing," NBER Working Paper No. 10915, 2004.

Rotemberg, Julio J., "Customer Anger at Price Increases, Changes in the Frequency of Price Adjustment and Monetary Policy," *Journal of Monetary Economics*, LII (2005), 829-852.

Rotemberg, Julio J. , "Expected Firm Altruism and Brand Extensions," Harvard Business School Working Paper, 2008.

Selten, Reinhard, and Axel Ockenfels, "An Experimental Solidarity Game," *Journal of Economic Behavior and Organization*, XXXIV (1998), 517-39.

Sobel, Joel, "The Timing of Sales," *The Review of Economic Studies*, LI (1984), 353-368.

Sobel, Joel, "Durable Goods Monopoly with Entry of New Consumers," *Econometrica*, LIX (1991), 1455-1485.

Tadelis, Steven, "What's in a Name? Reputation as a Tradeable Asset," *American Economic Review*, LXXXIX (1999), 548-563.

Weiss, Yoram, "Inflation and Price Adjustment: A Survey of Findings from Micro Data," in *Optimizing Pricing, Inflation, and the Cost of Price Adjustment*, E. Sheshinski and Y. Weiss, ed. (Cambridge, MA: MIT Press, 1993).

Wernerfelt, Birger, "Umbrella Branding as a Signal of New Product Quality: An Example of Signalling by Posting a Bond," *Rand Journal of Economics*, XIX (1988), 458-66.

Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen, "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *The Review of Economics and Statistics*, LXXXVI (2004), 514-533.

Table I
Check on Randomization Procedure

	Shallow Discount Condition	Deep Discount Condition	Difference	p-value
Days since last purchase (Recency)	646.04 (2.54)	647.92 (3.61)	-1.88 (4.42)	0.67
Number of units purchased (Frequency)	3.34 (0.02)	3.31 (0.03)	0.03 (0.04)	0.42
Average price of units purchased (Monetary Value)	133.26 (0.44)	132.69 (0.63)	0.57 (0.77)	0.46
Sample Size	36,815	18,232		

Table I reports the mean values of each historical purchasing measure (calculated separately for each condition). The statistics are calculated using purchases during the 8-year pre-test period, prior to the mailing date for the Test Catalog. Standard errors are in parentheses. The p-values denote the probability that the difference between the deep discount and shallow discount averages will be larger than the observed difference, under the null hypothesis that the true averages are identical.

Table II
Post-test Purchases Complete Sample

	Shallow Discount	Deep Discount	Difference
Revenue and Orders per Customer			
Revenue	$159.28 ($2.19)	$157.16 ($3.06)	-$2.11 ($3.78)
Orders: all post-test orders	1.05 (0.01)	1.04 (0.02)	-0.02 (0.02)
Orders: from the Test Catalog itself	0.022 (0.001)	0.036 (0.002)	0.014[**] (0.002)
Orders: from other catalogs	1.03 (0.01)	1.00 (0.02)	-0.03 (0.02)
% of customers with at least 1 post-test order	35.4% (0.3%)	35.0% (0.4%)	-0.4% (0.4%)
Number of customers	36,815	18,232	
Composition of Orders			
Average item price	$101.49 ($0.51)	$101.14 ($0.74)	- $0.35 ($0.90)
Average number of items per order	1.57 (0.01)	1.60 (0.01)	0.03 (0.02)

Table II reports the averages of each post-test sales measure for the respective samples.   All statistics are calculated using only purchases from the 28-month post-test period (after the Test Catalog was mailed).   The *orders*, *revenue* and *any post-test order* statistics are calculated using the 36,815 and 18,232 customers in the Shallow and Deep Discount conditions (respectively). The average item price and number of items per order are calculated using the 13,038 and 6,388 customers from each condition who placed at least one order during the post-test period. Standard errors are in parentheses.
[**]Significantly different from zero, $p < 0.01$.
[*]Significantly different from zero, $p < 0.05$.

Table III
Customers Who Paid a High Price for a Test Item
Within 3-Months of Receiving the Test Catalog

	Shallow Discount	Deep Discount	Difference
Revenue and Orders per Customer			
Revenue	$506.02 ($23.70)	$415.31 ($30.91)	-$90.71[*] ($40.15)
Orders: all post-test orders	3.27 (0.13)	2.78 (0.18)	-0.48[*] (0.22)
Orders: from the Test Catalog	0.10 (0.01)	0.12 (0.02)	0.02 (0.02)
Orders: from other catalogs	3.17 (0.13)	2.66 (0.17)	-0.50[*] (0.22)
% of customers with at least 1 post-test order	72.9% (1.5%)	66.0% (2.2%)	-6.9%[*] (2.6%)
Number of customers	933	459	
Composition of Orders			
Average item price	$97.61 ($1.98)	$97.82 ($2.87)	$0.21 ($3.53)
Average number of items per order	1.66 (0.06)	1.64 (0.06)	-0.02 (0.06)

Table III reports the averages of each post-test sales measure for the respective samples. The samples are restricted to customers who had paid a high price (above the shallow discount price) for a Test Item within 3-months of the Test Catalog mailing date. The *orders*, *revenue* and *any post-test order* statistics are calculated using all of the customers in the respective samples. The average item price and number of items per order are calculated using customers who placed at least one order during the post-test period. Standard errors are in parentheses.
[**]Significantly different from zero, p < 0.01.
[*]Significantly different from zero, p < 0.05.

Table IV
Poisson Regression: Impact of Deep Discounts on Post-test Orders

	Column 1	Column 2
	Customers With Prior Test Item Purchases	Customers Without Prior Test Item Purchases
Deep Discount	-0.078**	0.085
	(0.020)	(0.071)
Deep Discount * Time	0.026**	-0.017
	(0.007)	(0.012)
Time (since prior Test Item purchase)	-0.084**	-0.382**
	(0.006)	(0.007)
Recency (since any purchase)	-0.249**	
	(0.005)	
Frequency	0.620**	0.466**
	(0.005)	(0.012)
Monetary Value	0.038**	0.048**
	(0.010)	(0.014)
Intercept	0.926**	1.247**
	(0.051)	(0.081)
Log likelihood	-53,680	-30,308
Sample size	29,105	25,942

Table IV reports the coefficients from estimating Equation (2) on each sub-sample of customers. The dependent variable measures the number of orders made during the post-test period. Asymptotic standard errors are in parentheses.
**Significantly different from zero, $p < 0.01$.
*Significantly different from zero, $p < 0.05$.

Table V
Impact of Deep Discounts on Post-test Sales
By Timing of Prior Test Item Purchase

	Segment 1	Segment 2	Segment 3	Segment 4
	Under 250 days	250 to 500 days	500 to 750 days	Over 750 days
Deep Discount	-0.052**	-0.042	-0.005	0.079**
	(0.015)	(0.027)	(0.037)	(0.029)
Recency	-0.152**	-0.182**	-0.262**	-0.306**
	(0.009)	(0.012)	(0.017)	(0.011)
Frequency	0.623**	0.707**	0.769**	0.677**
	(0.010)	(0.016)	(0.024)	(0.018)
Monetary Value	-0.203**	-0.239**	-0.347**	-0.365**
	(0.016)	(0.026)	(0.033)	(0.025)
Intercept	1.724**	1.739**	2.397**	2.728**
	(0.085)	(0.132)	(0.170)	(0.136)
Log likelihood	-11,261	-6,364	-4,435	-8,459
Sample size	4,413	2,913	2,604	6,079

Table V reports the coefficients from estimating Equation (3) separately on the four customer segments.
Asymptotic standard errors are in parentheses.
**Significantly different from zero, $p < 0.01$.
*Significantly different from zero, $p < 0.05$.

Table VI
Impact of the Deep Discounts on
Posttest Orders from the Test Catalog and Other Catalogs

Timing of Prior Purchase	Column 1 Orders from the Test Catalog	Column 2 Orders from Other Catalogs	Column 3 All Posttest Orders
under 250 days	31.6%[**]	-6.5%[**]	-5.2%[**]
250 to 500 days	21.4%	-4.9%	-4.2%
500 to 750 days	84.3%[**]	-2.6%	-0.5%
over 750 days	61.2%[**]	6.6%[*]	7.9%[**]

Table VI reports the *Deep Discount* coefficient (β_1) from estimating Equation (3) separately on each customer segment and dependent variable. The results in Column 3 are also reported in Table V (and in Figure IV).
[**]Significantly different from zero, $p < 0.01$.
[*]Significantly different from zero, $p < 0.05$.

Table VII
The Role of the Price Paid in Earlier Transactions

	Past Price Segment		
	Low	Medium	High
Deep Discount	0.469	0.008	-0.133[**]
	(0.256)	(0.036)	(0.024)
Deep Discount * Time	-0.169[**]	-0.004	0.050[**]
	(0.063)	(0.013)	(0.009)
Time (since prior Test Item purchase)	0.045	-0.069[**]	-0.102[**]
	(0.055)	(0.010)	(0.007)
Recency (since any purchase)	-0.113[*]	-0.253[**]	-0.246[**]
	(0.056)	(0.008)	(0.006)
Frequency	0.489[**]	0.648[**]	0.606[**]
	(0.061)	(0.008)	(0.007)
Monetary Value	0.061	0.112[**]	0.008
	(0.073)	(0.017)	(0.013)
Intercept	0.056	0.477[**]	1.128[**]
	(0.037)	(0.090)	(0.068)
Log likelihood	-965	-21,449	-31,196
Sample size	463	12,633	16,009

Table VII reports the coefficients from estimating Equation (3) separately on the three customer segments. The Low segment includes customers who paid less than the deep discount price for a Test Item before the Test Catalog; customers in the Medium segment paid between the deep and shallow discount prices; customers in the High segment paid over the shallow discount price. The dependent variable in all three columns measures the number of units purchased during the post-test period. Asymptotic standard errors are in parentheses.
[**]Significantly different from zero, $p < 0.01$.
[*]Significantly different from zero, $p < 0.05$.

Table VIII

Additional Results: Revenue, Profits and Persistence of the Effect

	Column 1	Column 2	Column 3	Column 4
	Start of the Post-test Period	End of the Post-test Period	Revenue	Profits
Deep Discount	-0.139**	-0.127**	-65.83**	-48.42**
	(0.033)	(0.035)	(19.72)	(14.55)
Deep Discount * Time	0.044**	0.056**	21.30**	15.46**
	(0.013)	(0.013)	(6.74)	(4.98)
Time	-0.117**	-0.085**	-31.07**	-21.00**
	(0.010)	(0.011)	(6.22)	(4.59)
Recency	-0.249**	-0.242**	-78.24**	-59.14**
	(0.009)	(0.009)	(5.43)	(4.01)
Frequency	0.598**	0.616**	204.23**	150.84**
	(0.009)	(0.010)	(4.26)	(3.14)
Monetary Value	0.040**	0.025**	58.22**	45.61**
	(0.018)	(0.018)	(6.03)	(4.45)
Intercept	0.375**	0.495**	277.56**	190.79**
	(0.095)	(0.097)	(37.66)	(27.80)
Log likelihood	-19,674	-20,635		
Adjusted R^2			0.229	0.230
Sample size	16,009	16,009	16,009	16,009

The dependent variable in the four columns are: the number of units purchased in the first fourteen months of the post-test period (Column 1); the number of units purchased in the first fourteen months of the post-test period (Column 2); total revenue earned from each customer in the post-test period (Column 3); and total profit (calculated as revenue minus cost of goods sold) earned from each customer in the post-test period (Column 4). Columns 1 and 2 are estimated using Poisson regression (asymptotic standard errors are in parentheses). Columns 3 and 4 are estimated using OLS (standard errors are in parentheses). All of the models are estimated using the 16,009 customers who had previously paid a high price for a Test Item.
**Significantly different from zero, $p < 0.01$.
*Significantly different from zero, $p < 0.05$.

Table IX
Were Customers Delaying in Anticipation of Future Discounts?
Post-test Purchases at Different Discount Levels

Discount Threshold	Coefficients of Interest	
	(β_1)	(β_2)
	Deep Discount	Deep Discount * Time
Any discount level	-0.130** (0.018)	0.055** (0.007)
Discounts of at least 10%	-0.115** (0.020)	0.049** (0.008)
Discounts of at least 20%	-0.114** (0.020)	0.048** (0.008)
Discounts of at least 30%	-0.112** (0.021)	0.048** (0.008)
Discounts of at least 40%	-0.111** (0.021)	0.050** (0.008)
Discounts of at least 50%	-0.115** (0.021)	0.050** (0.008)
Discounts of at least 60%	-0.109** (0.024)	0.049** (0.009)
Discounts of at least 70%	-0.091** (0.034)	0.066** (0.013)

Each row in Table IX reports the β_1 and β_2 coefficients from re-estimating Equation (2) when the dependent variable counts post-test units sold under each discount threshold. In each of these eight models the samples are restricted to customers who had paid a high price (the sample size for each model is 16,009). Asymptotic standard errors are in parentheses.
[†]Significantly different from zero, $p < 0.10$.
[*]Significantly different from zero, $p < 0.05$.

Table X
Impact of Deep Discounts on the Number of Units
Purchased in the Post-test Period

	Column 1	Column 2
Deep Discount	-0.126[**]	-0.146[**]
	(0.053)	(0.026)
Deep Discount * Time	0.051[**]	0.051[**]
	(0.010)	(0.009)
Deep Discount * Maximum Past Discount	-0.022	
	(0.089)	
Deep Discount * 60% Past Discount		0.049
		(0.032)
Time	-0.097[**]	-0.102[**]
	(0.007)	(0.007)
Maximum Past Discount	1.004	
	(0.062)	
60% Past Discount		-0.018
		(0.019)
Recency	-0.239[**]	-0.246[**]
	(0.006)	(0.006)
Frequency	0.528[**]	0.607[**]
	(0.008)	(0.007)
Monetary Value	0.143[**]	0.008
	(0.015)	(0.013)
Intercept	0.098	1.134[**]
	(0.089)	(0.068)
Log likelihood	-31,024	-31,194
Sample size	16,009	16,009

Table X reports the coefficients from estimating Equation (4) on the customers who had previously paid a high price for a Test Item. Asymptotic standard errors are in parentheses.
[**]Significantly different from zero, $p < 0.01$.
[*]Significantly different from zero, $p < 0.05$.

Table XI
Impact of Deep Discounts on the Number of Units
Purchased in the Post-test Period

	Column 1	Column 2
	Non-Test Items	Test Items
Deep Discount	-0.152**	-0.047
	(0.021)	(0.039)
Deep Discount * Time	0.057**	0.046**
	(0.008)	(0.015)
Time (since prior Test Item purchase)	-0.153**	-0.176**
	(0.007)	(0.013)
Recency (since any purchase)	-0.196**	-0.180**
	(0.006)	(0.012)
Frequency	0.752**	0.514**
	(0.006)	(0.011)
Monetary Value	0.116**	-0.177**
	(0.012)	(0.021)
Intercept	0.483**	0.917**
	(0.062)	(0.111)
Log likelihood	-40,946	-17,817
Sample size	16,009	16,009

Both models are estimated using the 16,009 customers who had previously paid a high price
for a Test Item. In Column 1 the dependent variable measures the number of non-Test Items
purchased during the post-test period. In Column 2 the dependent variable measures the
number of Test Items purchased. Asymptotic standard errors are in parentheses.
**Significantly different from zero, $p < 0.01$.
*Significantly different from zero, $p < 0.05$.

Table XII
Replication Results
Impact of Sale Catalog on Post-Test Orders from the Apparel Catalog

	Column 1	Column 2	Column 3
	Customers Who Had Previously Paid More Than the Sale Price	Customers Who Had Paid a Price No Higher Than the Sale Price	Customers Who Had Not Previously Purchased a Sale Item
Received Sale Catalog	-0.024[**] (0.005)	0.012 (0.013)	0.016[**] (0.003)
Recency	-0.208[**] (0.002)	-0.177[**] (0.006)	-0.220[**] (0.001)
Frequency	0.624[**] (0.003)	0.673[**] (0.007)	0.541[**] (0.002)
Monetary Value	0.002[**] (0.011)	0.081[**] (0.028)	0.075[**] (0.005)
Intercept	0.478[**] (0.048)	0.047 (0.116)	0.511[**] (0.021)
Log likelihood	-74,498	-10,392	-362,018
Sample size	14,699	1,630	93,142

The dependent variable in all three models measures the number of orders made during the post-test period. Column 1 is estimated using the 14,699 customers who had previously paid a higher price for one of the items in the Sale catalog. Column 2 is estimated using the 1,630 customers who had previously paid a price equal to or less than the price in the Sale catalog. Column 3 is estimated using the 93,142 customers who had not previously purchased one of the items in the Sale catalog. Asymptotic standard errors are in parentheses.
[**]Significantly different from zero, $p < 0.01$.
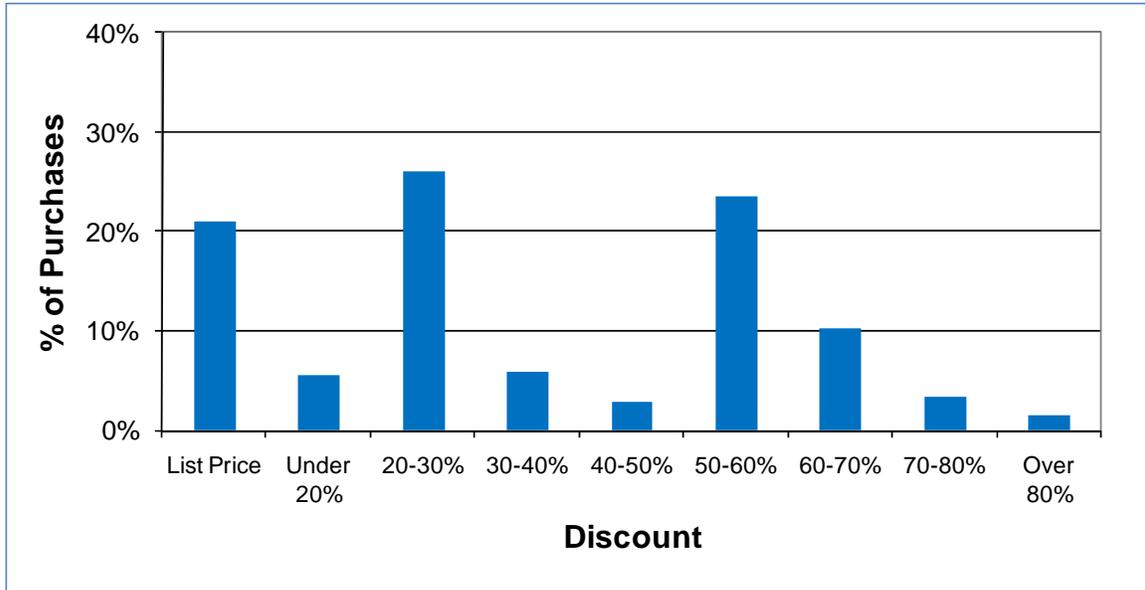[*]Significantly different from zero, $p < 0.05$.

Figure I reports a histogram of the discounts received in purchases in the two years before the Test Catalog was mailed. The histogram was calculated using a randomly selected sample of customers (including customers not involved in the study).
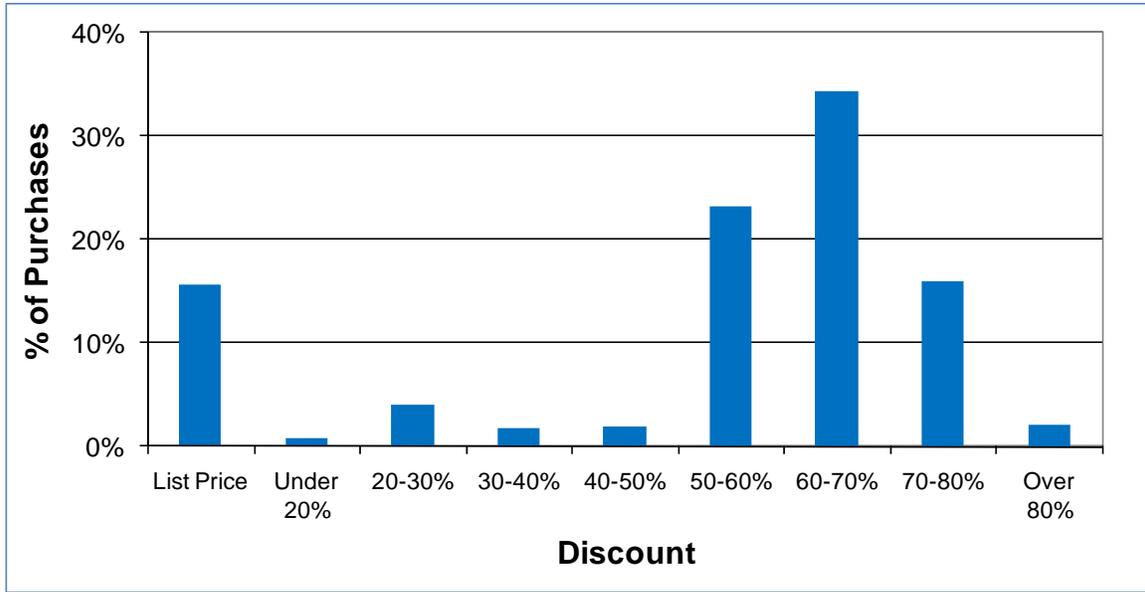
Figure I. Histogram of Pre-test Discounts

Figure II reports a histogram of the discounts received in the 28-month post-test period. The histogram was calculated using a randomly selected sample of customers (including customers not involved in the study).
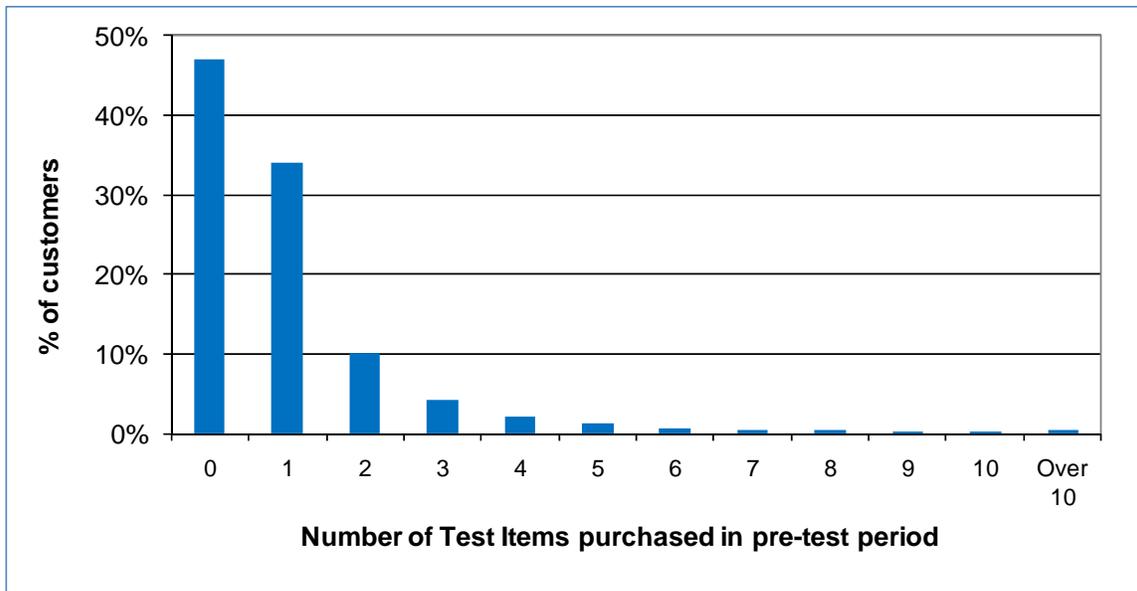
Figure II. Histogram of Post-test Discounts

Figure III presents a frequency distribution of the number of Test Items purchased in the pre-test period by all 55,047 customers involved in the first study. The unit of observation is an individual customer. The columns add to 100%.
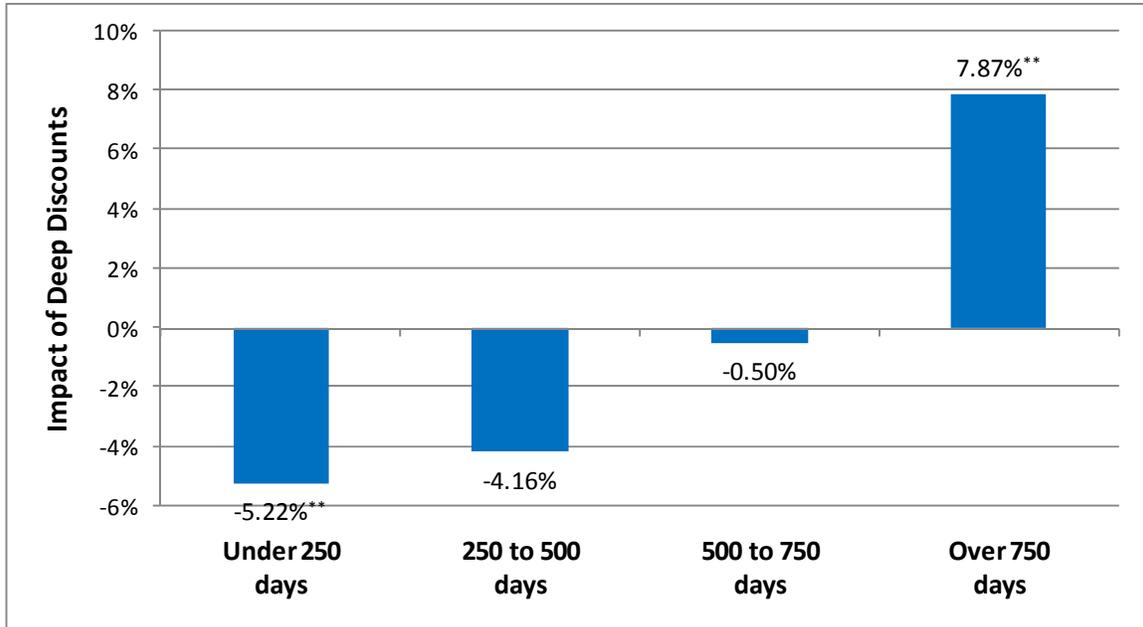
Figure III. Pre-test Purchases of Test Items

Figure IV reports the β_1 coefficients when estimating Equation (3) on the four customer segments described above. Detailed findings are provided in Table V.

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

**Figure IV. Impact of Deep Discounts on Post-test Sales
By Timing of Prior Test Item Purchase**