Measuring and Mitigating the Costs of Stockouts

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

Gavan J. Fitzsimons  
Fuqua School of Business, Duke University, 1 Towerview Drive, Durham, North Carolina 27708, gavan@duke.edu

Duncan Simester  
Sloan School of Management, Massachusetts Institute of Technology, Room E56-305, 38 Memorial Drive, Cambridge, Massachusetts 02142, simester@mit.edu

There is now an extensive theoretical literature investigating optimal inventory policies for retailers. Yet several recent reviews have recognized that these models are rarely applied in practice. One explanation for the paucity of practical applications is the difficulty of measuring how stockouts affect both current and future demand. In this paper, we report the findings of a large-scale field test that measures the short- and long-run opportunity cost of a stockout. The findings confirm that the adverse impact of a stockout extends to both other items in the current order as well as future orders.

We show how the findings can be used to provide input to inventory planning models and illustrate how failing to account for the long-run effects of a stockout will lead to suboptimal inventory decisions. We also demonstrate how the findings can be used in a customer lifetime value model. Finally, the study investigates the effectiveness of different responses that firms can offer to mitigate the cost of stockouts. There is considerable variation in the effectiveness of these responses. Offering discounts to encourage customers to backorder rather than cancel their orders is widely used in practice, but that was the least profitable of the responses that we evaluated. The findings have important implications for retailers considering the use of discounts as a response to stockouts.

Key words: inventory; long run; stockouts

History: Accepted by Candace A. Yano, operations and supply chain management; received December 9, 2004.

This paper was with the authors 8 months for 3 revisions.

1. Introduction

Understanding the cost of a stockout is critical to the implementation of any retail inventory model. Unless these costs are known, retailers cannot balance the costs (and risk) of holding inventory with the forgone profits when an item is out of stock. The relevant costs include both the lost profits from the immediate order because of cancellations, and the long-run costs if stockouts reduce the likelihood of future orders. While there is an extensive literature investigating optimal inventory policies, there has been surprisingly little work investigating how customers respond when an item is out of stock. In most instances, these short- and long-run costs are simply unknown. The absence of reliable estimates of these costs is one explanation for why sophisticated inventory models have not seen broad adoption by retailers. Zipkin (2000) concludes that the absence of these measures has prompted many firms to simply give up trying to implement sophisticated inventory models. Similarly, Raman (1999) argues that the absence of accurate estimates of stockout costs explains, at least in part, why management science models have rarely been applied to managing fashion inventories.

In this paper, we conduct a field test in a mail-order catalog to measure both the short- and long-run costs of a stockout to a catalog company. To evaluate the short-run cost, we compare how stockouts affect the final disposition of both the out-of-stock item and other items in the order. To evaluate the long-run costs, we tracked customers’ subsequent purchasing in a 13-month posttreatment period. We then show how the findings can be used to provide inputs to inventory planning and customer lifetime value models.

The challenge is not just to understand how costly a stockout is; we would also like to know how to respond to mitigate these costs when a stockout arises. After consulting with the catalog managers, we designed five different responses that the firm’s customer service representatives offered to customers
when an item was out of stock. The quasi-experimental design provides a comparison of the effectiveness of each response at preserving the current order and encouraging customers to repurchase in the future.

Not all of the findings in the study generalize readily to other retailers. In particular, the specific magnitudes of the short- and long-run costs may vary across retailers. However, this paper does offer several contributions that generalize to other direct marketing retailers (though perhaps not to traditional retail stores). The primary methodological contribution of the study is to confirm the feasibility of measuring the costs of a stockout and to illustrate how the findings can be used in inventory planning and customer lifetime value models.

There are also at least four substantive contributions that we expect to generalize to other markets. First, the findings confirm that the impact of stockouts extends beyond the short term to also affect long-run profits. We observe significant reductions in long-run demand among customers who experience a stockout. Second, we find that stockouts adversely impact not just the profits earned on the out-of-stock item, but also other items in that order. This potentially has important implications for inventory policies for multiproduct retailers. Third, the findings provide clear evidence that firms can mitigate the cost of stockouts by changing the explanations that their customer service representatives offer to customers when items are out of stock. There is considerable variation in the effectiveness of the five responses at both preserving the current order and encouraging future purchases. Surprisingly, offering discounts to encourage customers to backorder rather than cancel their orders was not profitable. Finally, we present a series of findings illustrating how the cost of stockouts varies according to the characteristics of the customer, item, and order. While the magnitude of these findings may vary across retailers, many of the directional conclusions are expected to generalize.

**Previous Theoretical Literature**

The stream of research most closely related to this paper is the recent literature investigating how the long-run impact of stockouts affects firms’ optimal inventory policies. This relationship was explicitly modeled by Hall and Porteus (2000), who present a model of a dynamic game between two firms. The model predicts that incorporating the opportunity cost of lost future business into inventory policies will raise inventory levels and lower stockout rates. They simulate their model to investigate its properties for a newsvendor model.

Gaur and Park (2006) build on the Hall and Porteus (2000) result by considering a model with asymmetric customer learning. When customers experience positive or negative service encounters, they update their expectations about future encounters. In the context of stockouts, ordering an item that is out of stock prompts customers to anticipate a higher probability of a stockout on future orders. Customers purchase based on these expectations and then update their expectations for future purchase occasions. Retailers, in turn, compete on stockout levels, which allows the model to decompose the effects of both customer learning and competition on inventory levels. While customers in the Hall and Porteus (2000) model only react to the last service encounter, in the Gaur and Park (2006) model, they are Bayesian updaters who remember their entire service history.

Other models of inventory and competition can be found in Dana (2001) and Dana and Petruzzi (2001). Dana (2001) studies competition between firms when customers cannot observe inventory levels. He shows that high prices may provide a credible signal of product availability. Dana and Petruzzi (2001) consider a newsvendor model in which consumer demand is a function of the service level. Like Hall and Porteus (2000), they show that endogenizing the effect of stockouts increases fill rates (and inventory levels).

The papers described above represent only a small sample of what is an extensive theoretical literature investigating both inventory planning and customer lifetime value. We will discuss this literature in greater detail toward the end of the paper, where we use inventory and customer lifetime models to illustrate the implications of the findings.

**Previous Empirical Literature**

We have not found any previous experimental studies conducted in actual retail settings that investigate either the cost of a stockout or the effectiveness of different marketing strategies for reducing these costs. However, there have been studies conducted in simulated retail settings and studies that compare historical data. Charlton and Ehrenberg (1976) conducted an experiment in which they visited the homes of detergent purchasers over a 24-week period, offering subjects choices of artificial brands. During the 22nd and 23rd weeks, subjects were informed that the leading market share brand in the study was temporarily unavailable, and in Week 24, the brand was reintroduced. Market shares and category sales returned to their preconstraint levels, with no apparent carry-over effects of the stockout. Motes and Castleberry (1985) performed a partial replication of this study using real potato chip brands. In their study, market shares did not return to preconstraint levels, although category sales did return to their original level. In an effort to clarify these discrepancies, Straughn (1991) examined the relationship between stockouts and market
share using weekly store-level scanner panel data (see also Campo et al. 2003). Her results showed a negligible short-term relationship between stockouts and market share but found a strong negative relationship between stockouts and long-term market share, particularly for large market share brands. In a series of laboratory studies, Fitzsimons (2000) found similar support for the potential long-term effects of stockouts. Consumers exposed to a stockout in a simulated Internet retailing environment showed dramatically higher rates of store switching than comparable consumers not exposed to a stockout.

There are also several studies that report findings from customer surveys. Schary and Christopher (1979) find that nearly half of all consumers who self-reported that an item was out of stock intended to make their purchase at another retail outlet, while approximately 20% reported they would not make any purchase. Schary and Christopher (1979) also asked consumers to rate the image of the store. Store image ratings were lower for consumers who had reported a stockout than for consumers who had not. Walter and Grabner (1975) employed a similar methodology to estimate the costs of single and repeated stockouts based on consumer responses to their survey. A more recent study (Emmelhainz et al. 1991) performed in-store interviews with consumers who experienced stockouts and indicated that 32% switched brands, 41% purchased a different size or variety, while 14% planned to go to another store.

2. Quasi-Experimental Design and Measures
The field test was conducted in a mail-order catalog that sells bedding and home accessories in the moderate to premium price range. Although for confidentiality reasons, we unable to identify the name of the catalog, at the end of this section, we provide a more detailed description of the products and characteristics of the customers and their orders. Initial discussions with catalog management revealed that the company tracked the proportion of items that were unavailable when customers ordered, but did not track the outcome of these orders. Management had considered a range of potential responses to stockouts, but had not systematically evaluated their effectiveness.

The research design included the use of nonequivalent control groups, pretest measures of historical purchasing behavior, and posttest tracking of the outcome of each order. We depict the design in Figure 1 using Cook and Campbell’s (1979) notation. Measurement activities are denoted by O, × denotes treatment activities, and the dashed line distinguishes between treatment groups. The treatment activities all occurred during the five-week treatment period. During this period, orders were received from a total of 22,921 customers, representing orders for more than 63,051 items (many orders include more than one item).

Customers who called to place an order during this period were assigned to one of six quasi-experimental conditions. If the customer ordered an item that was out of stock, the telephone service representative responded using one of five scripts (denoted by ×₁ through ×₅). The assignment of customers to these groups is discussed below. Customers who did not experience an out-of-stock situation did not receive any of the experimental treatments (this sample is denoted by the last row in Figure 1).

Measurements were obtained before and after the treatment period. The measurements prior to the treatment period are denoted by Oₓ and include the history of each customer’s transactions in the preceding 21 months. Unfortunately, the company did not have any data describing the incidence of out-of-stock items during this pretreatment period. In the 13-month posttreatment period, we collected data describing the subsequent orders placed by each of the 22,921 customers who ordered during the treatment period (denoted by Oᵧ).

The disposition of the orders received during the treatment period was evaluated by tracking the status of each order. Out-of-stock situations are often temporary, lasting only until new inventory arrives from the supplier. Therefore, items that were not in stock at the time of the order may be subsequently shipped if the customer backorders rather than cancels the item. This results in the following three possible final dispositions for an order:

(1) **Shipped not returned**: Shipped to the customer and not returned.

(2) **Shipped but returned**: Shipped to the customer and returned by the customer.

(3) **Cancelled**: The order for the item had been cancelled.

While exchanges of items are rare, when they occur, they are treated as **shipped but returned** for the original item and a new order for the replacement item.
Mitigating the Impact of a Stockout

During the treatment period, customers who ordered an item that was out of stock received one of five responses from the telephone service representative. The design of the five responses resulted from discussions between catalog management and the research team. These discussions revealed that industry practice includes both financial and nonfinancial responses to stockout situations. Financial responses offer customers compensation to offset any negative reactions resulting from the unavailability of an item, generally structured as a discount on the item in question or a discount on shipping and handling charges. Nonfinancial responses offer an explanation for why the item is unavailable. The following five responses compared in this study are best illustrated by the verbal scripts that the telephone service representatives used when an item was out of stock:

1. **Standard response**: “This item is out of stock.”
2. **Supplier problem**: “This item is out of stock because of a problem with our supplier.”
3. ** Extremely popular**: “This item is out of stock because it is extremely popular.”
4. **$5 off**: “This item is out of stock, but I can offer you $5 off of your shipping charges if you would like to wait for it.”
5. **10% off**: “This item is out of stock, but I can offer you a 10% discount on that item if you would like to wait for it.”

The first script represents the standard response (prior to this study) provided to customers who ordered an item that was unavailable. This condition was used as a benchmark against which to compare the four alternative responses. Of the remaining four scripts, two offer a financial incentive to the customer to backorder the unavailable item. In one condition, the incentive is a $5 discount on shipping, and in the other, it is a 10% discount on the purchase price of that item. The average price of an item is approximately $51, and so the dollar magnitude of the incentive was expected to be similar on average in these two conditions. The other two scripts, which we label supplier problem and extremely popular, offer explanations for why the item was unavailable and do not provide any financial incentive to encourage customers to place the item on backorder.

Telephone service representatives were trained in the use of the scripts approximately three weeks prior to the treatment period. Training included implementation of procedures to ensure that all demand was measured, whether or not it resulted in an order. For example, if a customer asked for an item that was unavailable and either cancelled the order or substituted another item, telephone service representatives were trained to record the demand for the initial item and then designate it as cancelled where appropriate. Adherence to the procedures during the treatment period was monitored by the call center supervisors, and during site visits, by the research team.

Assignment of Customers to the Treatment Conditions

When customers call the catalog’s toll-free telephone number, they are randomly assigned to an available telephone service representative. To reduce the potential for confounds, each telephone representative was trained to use just one of the scripts (when an item was unavailable). Assignment of the representatives to the scripts was undertaken by catalog management and was apparently random. The study included 61 representatives, although 3 of these representatives handled a negligible number of calls for the catalog during the treatment period (the representatives also take calls for other catalogs owned by this firm). Among the remaining 58 representatives, groups of 10, 12, 9, 16, and 11 representatives were assigned to Script 1 through Script 5 (respectively).

To investigate whether differences in the characteristics of the representatives may have affected the findings, we compared the orders for customers who did not experience a stockout. Recall that customers who did not experience a stockout did not receive any of the experimental treatments, and so any systematic differences in the outcome of these orders can be attributed to differences specific to the representatives rather than the experimental treatments. Univariate analysis yielded no significant differences in the size of these orders between the five groups of representatives. In our multivariate analysis, we include controls that account for any potential differences between these representatives (operators).

Characteristics of the Customers and Product Categories

To assist in interpreting the results, it is helpful to provide more information about the customers and the products. The items are almost all durables (rather than consumables), and repeat purchases of identical items are uncommon. There are strong complementary relationships between some of the products, including, for example, sheets and matching pillowcases, or bath towels and matching hand towels. Almost half of the items (46.8%) are private label items that are only available through this company’s catalog or Internet site. The remaining items are well-known national brands that are available at other retailers; this distinction will later prove helpful in interpreting the findings. More detailed descriptions of the product categories are reported in the online appendix on the Management Science website (http://mansci.pubs.informs.org/ecompanion.html).
Analysis of the historical transaction data revealed that of the 22,291 customers who purchase during the treatment period, 10,256 (46%) had also purchased in the 21-month pretreatment period. The remaining 12,035 (54%) had not placed an order in the pretreatment period. Among the 46% of customers for whom we have a prior purchase history, the average purchase was 2.18 items at an average cost of $101.28. On average, the most recent prior purchase had occurred 237 days before the start of the treatment period.

3. The Short-Run Impact on the Current Order
Approximately 21.9% of the items ordered during the treatment period were out of stock at the time of the order. Because orders generally included more than one item, an average of 31.6% of orders received during the treatment period included at least one item that was unavailable. While this incidence of stockouts may appear surprisingly high, the firm’s managers reported that these proportions are consistent with historical stockout levels at this catalog. As our results on the optimal inventory policy for this firm will later show, this fill rate is consistent with the firm not giving any weight to the cost of a stockout in its inventory policy. More generally, it highlights the relative observability of obsolescence and stockout costs. Disposing of obsolete items at clearance prices or through a jobber is an event that is much more easily observed than the somewhat hidden costs of a stockout. Recall that the firm had not previously measured the cost of a stockout. Observability brings accountability, so it is natural that obsolescence costs receive greater weight than (unmeasured) stockout costs in a firm’s inventory policy.

In Table 1, we summarize the outcome of each order and compare how the outcome varied across items that were in stock and out of stock at the time of the order. The catalog only earned revenue on items that were in stock and out of stock at the time of the order and compare how the outcome varied across

<table>
<thead>
<tr>
<th>Category</th>
<th>In-stock items</th>
<th>Out-of-stock items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancelled (%)</td>
<td>4</td>
<td>32</td>
</tr>
<tr>
<td>Shipped but returned (%)</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Shipped not returned (%)</td>
<td>86</td>
<td>62</td>
</tr>
<tr>
<td>Sample size</td>
<td>49,235</td>
<td>13,816</td>
</tr>
</tbody>
</table>

Note. The dispositions are mutually exclusive and exhaustive so that each column totals 100%.

The difference in the conversion rate for in-stock and out-of-stock items is stark. The catalog earned revenue on 86% of the items ordered that were in stock at the time of the order. The conversion rate dropped to 62% if the item was out of stock at the time of the order. If the conversion rate in the treatment period had been 86% on all items, the increase in revenue would have been approximately $180,000. For completeness, we also report the conversion rates by product category in Anderson et al. (2006).

The comparisons in Table 1 do not account for possible substitution of an alternative item when an item is out of stock. Further investigation revealed that the level of substitution was negligible. The telephone service representatives and catalog management did not find this result surprising. They indicated that customers of this catalog order specific items and reported that customers are rarely willing to substitute for an alternative item. We caution that this finding alone should not be used as justification for ignoring substitution when measuring the impact of stockouts. The finding could be specific to the product categories offered by this company and may not generalize to other categories.

Proportion of Out-of-Stock Items in the Order
The findings in Table 1 confirm that a stockout on an item affects the conversion rate for that item. We were also interested in whether a stockout affects the conversion rate of other (nonfocal) items in the order. We investigate this issue, for each item ordered during the treatment period, we calculated the average conversion rate for other items in the order. Under the null hypothesis that stockouts do not affect the conversion rate of other items, this average should be unaffected by the availability of the focal item.

The data support rejection of the null hypothesis. If the focal item is in stock, the conversion rate on other items in the order is 83.1%. In contrast, if the focal item is out of stock, the conversion rate on other items falls to 71.2%. The difference in these two conversion rates is significant ($p < 0.01$). We conclude that the adverse impact of a stockout extends beyond the focal item to also affect other items in the order. Customers are more likely to cancel items that are available if they are not able to get other items that they want. This may in part reflect the presence of complementary products in the catalog, such as matching sheets and pillowcases. If the sheet is unavailable, then customers have less need for the pillowcases.

Mitigating the Short-Run Impact of Stockouts
The data also allow us to compare how the different customer service responses affected the final disposition of items that were out of stock at the time of the order. In Table 2, we summarize the final disposition under each of the five experimental conditions.
The extremely popular and 10% off responses were the most effective at preserving orders for out-of-stock items. While there were no significant differences between these two scripts, both scripts had significantly (p < 0.01) lower cancellation rates and significantly (p < 0.01) higher conversion rates than each of the other responses.

The effectiveness of the extremely popular claim is perhaps surprising as the claim represents “cheap talk”; there is no way for customers to evaluate this claim. The finding is consistent with the argument that scarcity reinforces customers’ perceptions that an item is fashionable. The finding may also support Wernerfelt’s (1995) argument that decision makers who are unsure of their own (absolute) preferences use information about other customers’ preferences to guide their purchasing decisions. In particular, learning that a product is popular may prompt a customer to backorder the item if he or she believes that those preferences are similar to other customers’.

The evidence that offering a 10% discount is more successful at retaining the order than offering a $5 discount is also interesting. Recall that the average price of an item is approximately $51, so the dollar magnitude of the incentive was expected to be similar on average in these two conditions. We might expect that the 10% off condition would have performed relatively better than the $5 discount when the price of the items was higher. Further analysis reveals that the effectiveness of the 10% off response increases relative to the $5 off response when the price is higher. However, the 10% off response is more effective even when the resulting discount is less than $5.

To further investigate why customers responded differently to the different customer service scripts, we conducted a post-test survey of 101 students from a major university. Details of the survey are reported in the online appendix. This survey data suggested that the difference in conversion rates across the five scripts was because of different expectations about how quickly (rather than whether) customers would receive backordered items.

We caution that we do not intend these results to be interpreted as a recommendation that retailers mislead their customers by claiming that an item is extremely popular when it is not. Rather, the findings indicate that if an item is out of stock because demand exceeded the firm’s expectations, then simply explaining why it is out of stock is an effective strategy for preserving the order. The 10% off condition does offer the advantage that it can be used in any situation without misleading customers. However, it suffers from the disadvantage that it is costly. Because of this additional cost, the 10% off and $5 off responses were the least profitable of the five responses, yielding an average profit of just $15.97 and $14.28 on out-of-stock items, respectively.1 This compares with $18.18 in the control and $19.10 and $20.51 in the supplier problem and extremely popular conditions. The increase in the conversion rate that results from the financial incentives was insufficient to outweigh the cost of paying the incentives to every customer who backordered the item (including customers who would have backordered the item anyway). Of course, this cost would be lessened if the customer service representatives were successfully trained to offer the incentives only to customers who plan to cancel their order.

Anticipated Delivery Delay
There is an additional source of data that allows us to further investigate whether the difference in conversion rates across the five scripts was because of different expectations about how quickly customers would receive backordered items. The computer system used by the telephone service representatives estimates the anticipated delay before an out-of-stock item is available to ship. In particular, the system calculates the difference between the current (order) date and the date that the next purchase order is expected to arrive from the supplier. Discussions with catalog management and monitoring of actual telephone calls revealed that this information is almost always provided to customers when an item is out of stock. Normally, this information is updated dynamically and is not stored. However, we were able to download this inventory data for a sample of approximately 30% of the orders.

The overall relationship between the anticipated delay and the conversion rate is summarized in Table 3 (using the subsample of available data). The outcome for in-stock items is included as a basis for comparison. Interestingly, even when the anticipated delay for an item is short (1–5 days), there is

---
1. The average profit was derived by calculating the total profit earned from out-of-stock items in each condition (less any discount) and dividing by the number of out-of-stock orders in that condition.
an immediate decrease in the conversion rate and a corresponding increase in the cancellation rate. The conversion and cancellation rates are fairly constant for anticipated delays of up to three weeks. However, after 20 days, there is a further decrease in the conversion rate and a corresponding increase in cancellations. One explanation for the increase in cancellations after 30 days is that once the delay exceeds this period, direct mail firms are required by the Federal Trade Commission to send customers written notification inviting them to cancel their orders.

A simple logit model of the conversion rate, with anticipated delay as the only independent variable, confirms that there is a significant negative association (p < 0.0001) between the conversion rate and the anticipated delay. However, while the average anticipated delay was the same for all five scripts (as we would expect), the relationship between the conversion rate and the anticipated delay varies across the five scripts. The relationship is (significantly) strongest in the control condition and weakest in the extremely popular and 10% off conditions.

One interpretation of these results is that the extremely popular response served as a partial substitute for the anticipated delay information provided by the customer service representatives. This interpretation is consistent with our earlier survey evidence that respondents in this condition reported that they anticipated shorter waits for the out-of-stock item to be delivered. While this evidence supports this interpretation, we hesitate to describe it as conclusive.

### Multivariate Analysis

The findings in the previous tables have focused on a univariate analysis of the data. We can also use a multivariate approach to investigate how several different factors jointly moderate the impact of a stockout on the conversion rate. In particular, we used a logit model to estimate whether the final disposition of each item ordered during the treatment period was shipped and not returned (the conversion probability). Additional details are available in the online appendix. As expected, we see that if an item is out of stock at the time of the order, there is a significant drop in the conversion rate across all five response conditions. The effect is larger if the item has a higher price but smaller if the item is a store brand. Recall that store brands are exclusive to the store, so evidence that customers are less likely to cancel store brands is consistent with customers recognizing that they will not be able to find the merchandise elsewhere. We see little evidence that the impact of a stockout varies across customers. We might have expected that customers who had made recent prior purchases (during the pretreatment period) would respond differently to a stockout than customers without recent prior purchases. However, the data do not support this prediction.

The remaining results are consistent with our earlier findings. The number of other items in the order that are out of stock plays an important moderating role. Customers are more likely to cancel an item if they are unable to get other items that they want. In contrast, they are less likely to cancel when more of the other items are available. Furthermore, when an item is out of stock, we see that the drop in the conversion rate is smallest in the extremely popular and 10% off conditions (and significantly different than in the other conditions).

### 4. The Long-Run Impact of a Stockout

While many of the actions that firms take affect their demand and profits in subsequent periods, the vast majority of studies limit attention to the immediate outcome. This at least in part reflects the difficulty of collecting data to measure long-run effects. However, there is growing interest in the long-run impact of firms’ actions, with several recent examples of studies that have sought to measure long-run effects (see, for example, Mela et al. 1997, Mela et al. 1998, Pauwels et al. 2002, Anderson and Simester 2004, and Simester et al. 2006). Many of these studies have focused on the impact of prices and price promotions.

In this study, we evaluate the long-run impact of stockouts by investigating how stockouts affected demand in the 13-month posttreatment period. Before presenting the findings, it is helpful to note an important difference in the unit of analysis. In the previous analysis, the unit of analysis was “an item ordered by a customer” during the treatment period, so that customers who ordered multiple items yielded multiple observations. The unit of analysis for the long-run analysis is “a customer.”

#### Table 3 Conversion Rate by Anticipated Delay

<table>
<thead>
<tr>
<th>Delay</th>
<th>Conversion rate (%)</th>
<th>Cancellation rate (%)</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>In stock</td>
<td>86</td>
<td>4</td>
<td>16,853</td>
</tr>
<tr>
<td>1–5 days</td>
<td>74</td>
<td>22</td>
<td>324</td>
</tr>
<tr>
<td>6–10 days</td>
<td>75</td>
<td>18</td>
<td>608</td>
</tr>
<tr>
<td>11–15 days</td>
<td>73</td>
<td>22</td>
<td>609</td>
</tr>
<tr>
<td>16–20 days</td>
<td>72</td>
<td>23</td>
<td>657</td>
</tr>
<tr>
<td>21–25 days</td>
<td>66</td>
<td>29</td>
<td>510</td>
</tr>
<tr>
<td>26–30 days</td>
<td>64</td>
<td>34</td>
<td>441</td>
</tr>
<tr>
<td>31–35 days</td>
<td>62</td>
<td>35</td>
<td>344</td>
</tr>
<tr>
<td>36–40 days</td>
<td>58</td>
<td>35</td>
<td>153</td>
</tr>
<tr>
<td>More than 40 days</td>
<td>54</td>
<td>41</td>
<td>727</td>
</tr>
<tr>
<td>Total</td>
<td>226</td>
<td></td>
<td>21,226</td>
</tr>
</tbody>
</table>
This change introduces a potential confounding factor: Customers who ordered more items in the treatment period were more likely to experience a stockout, and these customers were also more likely to order in the posttreatment period (they are the catalog’s best customers). For this reason, simply comparing posttreatment demand based on whether a customer experienced a stockout in the treatment period could lead to the erroneous conclusion that stockouts increase demand.

Fortunately, the historical purchasing data from the 21-month pretreatment period provide an opportunity to control for differences in these customer characteristics. In particular, we use the logs of the recency, frequency, and monetary value (RFM) of customers’ prior purchases to control for individual customer characteristics. These RFM measures are well-established metrics for segmenting customers in this industry and provide natural candidates for control variables. To ensure that we had historical data for every customer, when calculating the frequency and monetary value measures, we considered all orders prior to the posttreatment period. For the recency measure, we omitted orders during the treatment period, as every customer purchased during this period (including these purchases would have led to almost no variance in the recency measure).

For each customer, we calculated four dependent measures of posttreatment demand: the number of units ordered; total revenue; the conversion rate; and whether a customer made any purchase during the posttreatment period. The varying characteristics of these dependent measures call for different estimation models. In particular, the number of units purchased during the posttreatment period is a count measure that can be expected to follow a Poisson distribution, so we use Poisson regression to estimate this model. The total revenue model is estimated using a linear regression model, the conversion rate is estimated using a linear probability model, and any purchase is estimated using a logit model. The models use the full sample of 22,921 customers, with the exception of the conversion rate model, which is estimated using only those customers who purchased during the posttreatment period. For all four models, we use the following same vector of independent variables:

$$\beta X = \alpha + \beta_1 \ln(\text{Recency}_i) + \beta_2 \ln(\text{Frequency}_i)$$
$$+ \beta_3 \ln(\text{Monetary Value}_i)$$
$$+ \beta_4 \text{Percent Stockouts}_i. \quad (1)$$

The $\alpha$ and $\beta_k$ terms are estimated parameters, while the percent stockouts variable measures the percentage of items ordered by the customer during the treatment period that was out of stock at the time of the order. This measure is essentially uncorrelated with the number of items ordered (further controlling for customer differences). Under this specification, the $\beta_4$ coefficient measures the relationship between stockouts in the treatment period and demand in the posttreatment period.

The findings are reported in Table 4. They reveal that stockouts during the treatment period had an adverse impact on all four dependent measures of posttreatment demand. In particular, customers who experienced stockouts on a higher proportion of items during the treatment period were less likely to place an order in the posttreatment period, they tended to order fewer items, their total revenue was lower, and their conversion rate was lower. The effect is large: If all the items ordered by a customer during the treatment period were out of stock, posttreatment demand was approximately 22% lower than if none of the items was out of stock.

We caution that because the conversion rate model only includes customers who placed an order during the posttreatment period, the model is subject to a potential selection bias. In particular, we do not observe a conversion rate for customers who did not order, and the characteristics of these customers are likely to be different from those who did order.

In the findings reported in Table 4, we aggregated demand across the entire 13-month posttreatment period. By dividing the posttreatment period into subperiods, we can compare the impact of the deep discounts at the start and end of the posttreatment period. In particular, we divided the 13 months into two 6.5-month subperiods and re-estimated the
Mitigating the Long-Run Cost of Stockouts

By slightly modifying the model, we can also investigate whether the five customer service responses had any effect on the long-run impact of the stockouts. We modified the four models by introducing the following additional independent variables:

\[ \beta X = \beta_1 \ln(\text{Recency}_i) + \beta_2 \ln(\text{Frequency}_i) \]

\[ + \beta_3 \ln(\text{Monetary Value}_i) + \sum_{j=1}^{5} \alpha_j \text{Condition } j_i \]

\[ + \sum_{j=1}^{5} \eta_j \text{Condition } j_i \cdot \text{Percent Stockouts}, \]

The model has five intercepts (\(\alpha_j\)) and five interaction terms identifying the impact of stockouts under each of the five conditions. The condition \(J\) terms are the binary variables identifying which experimental condition a customer was in during the treatment period. Under this specification, the \(\eta_j\) coefficients estimate the relationship between stockouts in the treatment period and orders in the posttreatment period for customers in each of the five conditions.

The coefficients are reported in Table 6 (for ease of presentation, we omit the five intercept terms). The findings reveal that the long-run impact of the stockouts varied across the five experimental conditions but were generally consistent across the four models. The most favorable long-run findings were observed in the control and extremely popular conditions. For customers in these two conditions, the incidence of stockouts in the treatment period did not significantly impact demand during the posttreatment period. We should not conclude from the null results for these conditions that there was no long-run effect under these responses. Rather, it appears that the lack of statistical significance is largely because of the loss of statistical power when measuring the effects separately for each condition. For example, in the number of units model, when we aggregate the findings in the control and extremely popular conditions, we find that the joint effect is negative and significant (\(p < 0.05\)).

An adverse impact was also observed in the supplier problem condition, but only in the number of units model. Among customers in this condition, posttreatment demand fell by 14% if all the items they ordered during the treatment period were out of stock.

Significant negative outcomes were observed in the two conditions in which customers were offered discounts to encourage them to backorder rather than cancel. In the \$5 off condition, customers with no stockouts in the treatment period ordered 24% more units in the posttreatment period than customers with...
100% stockouts. This effect increased to more than 50% in the 10% off condition.

It is natural to ask why offering discounts to customers who experienced stockouts led to lower long-run demand. The attribution literature (e.g., Weiner 1980) offers a possible explanation. Researchers have found that when customers are offered a price promotion to induce them to purchase, they may attribute the purchase to the financial incentive and not the performance of the firm or product (see Blattberg and Neslin 1990, p. 29). These customers are less likely to make a repeat purchase than customers who purchased without a price promotion.

Alternatively, offering discounts to encourage customers to backorder may introduce concerns about the fairness of the original prices. In particular, customers may infer that if the catalog is still able to earn a profit after discounting the items, then it was earning excessive profits when charging the original prices. This explanation is motivated in part by the principle of dual entitlement proposed by Kahneman et al. (1986a, b). They argue that customers have perceived fairness levels for both firm profits and retail prices. While firms are entitled to earn a fair profit, customers are also entitled to a fair price. Deviations from a fair price can be justified only by the firm’s need to maintain a fair profit. Under this argument, it is fair for retailers to raise the price of snow shovels if the wholesale price increases, but it is not fair to do so if a snowstorm leads to excess demand. By revealing that it can earn a fair profit at the discounted prices, the firm may have implicitly revealed that it was earning an unfair profit at the original prices.

We caution that the findings do not establish that financial incentives at this firm will always lead to an adverse long-run outcome. It is possible that if the firm had framed the incentives differently, or had varied the size of the incentives, then the long-run outcome may have been different.

5. Implications for Inventory Planning and Customer Lifetime Value

The objective of a firm’s inventory policy is to increase profits, not revenue. In particular, optimal inventory policies trade off the cost of obsolescence with the opportunity cost of a stockout. This obviously requires a measure of the opportunity cost of a stockout. In this section, we show how to use the findings to calculate this opportunity cost. We do so in two stages, beginning with the short-run and followed by the long-run opportunity cost. We then illustrate how these opportunity costs can be used as inputs in both inventory planning and customer lifetime value models.

<table>
<thead>
<tr>
<th>Table 7 The Opportunity Cost of a Stockout</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-stock profit ($)</td>
</tr>
<tr>
<td>Direct short-run cost</td>
</tr>
<tr>
<td>Indirect short-run cost</td>
</tr>
<tr>
<td>Total short-run cost</td>
</tr>
<tr>
<td>Long-run cost</td>
</tr>
<tr>
<td>Overall cost</td>
</tr>
</tbody>
</table>

*p<0.05, **p<0.01.

The Short-Run Opportunity Cost of a Stockout

The short-run opportunity cost of a stockout includes two components: (1) the direct cost on the focal item and (2) the indirect cost on other items in the order. Recall that the conversion rate for an item depends not just on whether that item is in stock, but also on whether other items in the order are available. Therefore, if an item is out of stock, it has a direct impact on the profits earned from the order for that item and an indirect effect on the profits earned from other items in the order.

Calculating the direct short-run cost of a stockout is straightforward. We calculated the actual profit earned on each item ordered, representing the difference between the price paid (less any discount) and the cost of goods sold. For items that were cancelled or returned, the profit was zero. We then calculated the average profit for items that were in stock at the time of the order and compared it with the average for items that were out of stock. The results are summarized in Table 7, where, for the sake of brevity, we aggregate across the five response conditions.

To calculate the total short-run opportunity cost (including the direct and indirect effects), we use a “representative item” approach. For each customer, we randomly select one of the items that the customer ordered during the treatment period. Depending on whether this item was in stock, we assign the entire order to an “in-stock” or “out-of-stock” subsample. Note that this approach assigns a stockout status to the entire order rather than to just a single item. We then compare the average profits earned from orders in the two subsamples (see Table 7) and interpret the difference as the change in the short-run profitability of an order when a representative item is out of stock.

The findings reveal that the total short-run opportunity cost of a stockout averages $13.13 or 19%. Given that the firm only expects to earn an average of $25.96 from an item when it is in stock, this opportunity cost is relatively large. Moreover, the findings highlight the importance of considering both the direct and indirect effects. The indirect effect of a stockout on other nonfocal items in the order ($4.37) comprises approximately one-third of the total short-run cost ($13.13).
Long-Run Opportunity Cost of a Stockout
Recall that the unit of observation in the long-run analysis is a customer rather than an item. For this reason, calculating the long-run opportunity cost of a stockout again requires the representative item approach. For each customer, we randomly select one of the items that the customer ordered during the treatment period. Depending on whether this item was in stock, we assign the customer to an “in-stock sample” or “out-of-stock sample” and then compare the average profits across the two samples. The findings are also reported in Table 7. They reveal that, on average (across the five response conditions), long-run profits fell by $9.56 when a representative item was out of stock during the treatment period. Note that our estimates imply that the percentage loss in long-run profits is similar in magnitude to the percentage loss in short-run profits (19% versus 17%).

We caution that this estimate of the long-run opportunity cost of a stockout only considers the 13-month posttreatment period. The evidence in Table 5 that the adverse effects of a stockout persist throughout the 13-month posttreatment period suggests that the effects also extend into future periods. In this respect, the long-run opportunity cost estimated here represents a conservative estimate of the true cost.

Inputs to an Optimal Inventory Model
To illustrate how the short- and long-run costs of a stockout can be used as inputs in an optimal inventory model, we focus on the classic news-vendor problem (Porteus 2002). It is well established that the following optimal stockout rate can be written as (see, for example, Ravindran et al. 1987, p. 355):

\[
\text{Stockout Rate} = \frac{\text{Cost of Goods Sold} - \text{Salvage Value}}{\text{Price} + \text{Opportunity Cost} - \text{Salvage Value}}. \tag{3}
\]

For the average item sold by this retailer, it is straightforward to calculate the average cost of goods sold ($22.34), the average price ($51.06), and the average salvage value ($15.17).3 Substituting into Equation (3) yields the optimal stockout rate. In Table 8, we show how this rate varies according to which stockout costs the firm considers, including the benchmark of zero (ignoring the opportunity cost of a stockout altogether).

As we would expect, there is a sharp drop in the optimal stockout rate when weight is given to the opportunity cost of a stockout. Recall that the actual stockout rate at the firm during the treatment period was 21.9%. This analysis may help to explain why this rate was so high. The firm had not previously measured the opportunity cost of a stockout and was apparently not giving any weight to these costs in its inventory planning.

### Table 8: The Opportunity Cost of a Stockout

<table>
<thead>
<tr>
<th>Opportunity cost</th>
<th>Actual cost ($)</th>
<th>Optimal stockout rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>0.00</td>
<td>20.0</td>
</tr>
<tr>
<td>Direct short-run cost</td>
<td>8.76</td>
<td>16.1</td>
</tr>
<tr>
<td>Total short-run cost</td>
<td>13.13</td>
<td>14.6</td>
</tr>
<tr>
<td>Total short and long-run cost</td>
<td>22.69</td>
<td>12.2</td>
</tr>
</tbody>
</table>

Note. The table reports the results of calculating the optimal stockout rate using Equation (3).

### Inputs to a Customer Lifetime Value Model
There is a growing literature investigating how to estimate a customer’s lifetime value. To illustrate how our findings can be used to calculate how a stockout affects that value, we used the well-known Schmittlein et al. (1987) model.4 Under this model, the probability that a customer is still active (alive) given the past (observed) purchase history is given by

\[
\text{Pr(Active)} = \frac{1}{1 + (\mu/(\mu + \lambda))(e^{(\mu + \lambda)(T - t)} - 1)}, \tag{4}
\]

where \(\lambda\) is the (average) Poisson purchase rate for a customer who is still active, \(\mu\) is the death rate, \(t\) is the time since the last purchase, and \(T\) is the time since the first purchase. Our findings provide a measure of the buying rate and show how this is affected by a stockout. In particular, we used our representative item approach to assign customers to “in-stock” and “out-of-stock” subsamples and then calculated the average \(\lambda, t,\) and \(T\) for each subsample (as of the end of the posttreatment period). Unfortunately, the “death rate” \(\mu\) is not known, and because we cannot observe whether (or when) customers die, it is not easily estimated. In the absence of a known death rate, we use Equation (4) to calculate the probability that a customer is still alive for different values of \(\mu\). The findings are reported in Table 9.

The difference in the probability that a customer is alive across the two subsamples can be interpreted as the increased probability that a customer dies as a result of a stockout on a representative item. The increased risk of death is approximately 2%. Reassuringly, this probability is relatively stable across different values of the unknown parameter \(\mu\).

---

3 The company disposes of obsolete items through “clearance sales” and jobbers. They receive approximately 40% of the retail price in a clearance sale and just 12% from jobbers. The firm estimated that these two channels each receive a similar amount of volume, so we averaged these two salvage rates to arrive at the $15.17 average salvage value.

4 Recent extensions to this model include Ho et al. (2006) who show how to incorporate customer satisfaction into the model, and Fader et al. (2005) who propose an alternative specification that is easier to compute.
6. Discussion and Conclusions

We have reported findings from a large-scale field study conducted with a national mail-order catalog to identify the financial cost of a stockout and the effectiveness of different customer service responses to stockouts. The findings confirm that if an item is out of stock when a customer orders, the profitability of that order drops sharply for two reasons. The first effect is a direct effect on the focal item. Customers are more likely to cancel the out-of-stock item, so the firm is less likely to earn revenue. The second effect is an indirect effect on other items in the order. We find that a stockout on one item increases the probability of customers canceling other items in that order. This indirect effect is large, representing almost 33% of the total short-run opportunity cost. The size of this indirect cost may in part reflect the complementary nature of many of the products sold by this firm, but may also reflect shipping costs and other fixed costs associated with completing a transaction. In the face of these fixed costs, customers may have a preference for one-stop shopping, so that the unavailability of one item in a basket prompts them to search elsewhere for the entire basket.

The findings also confirm that stockouts can adversely impact long-run demand. Customers who experienced a stockout were less likely to place a subsequent order, ordered fewer items, spent less (revenue), and had a lower subsequent conversion rate. The long-run impact is also large—almost as large in magnitude as the short-run cost.

To investigate how firms can mitigate these costs, we compared the effectiveness of five different responses that customer service representatives often offer to customers when they order an item that is out of stock. Stating that the item is “extremely popular” was the most effective response for encouraging customers to backorder rather than cancel the item. This response was also the most effective at encouraging customers who experienced a stockout to reorder. The findings indicate that firms should use caution when offering discounts to encourage customers to backorder rather than cancel items that are out of stock. In the two conditions in which customers were offered discounts, stockouts had the most pronounced negative impact on long-run demand. These discounts appeared to aggravate rather than mitigate the costs of stockouts.

It is appropriate to consider the extent to which the findings from this study will generalize to other markets. As we acknowledged in the introduction to this paper, we do not believe that the specific cost estimates are directly generalizable. The magnitudes of these costs will almost certainly vary across firms and markets. However, there are several conclusions that we do expect will generalize. This includes the evidence that stockouts affect other items in the current order, affect future demand, and can be mitigated by varying the response that the customer service representatives offer.

The paper also makes a methodological contribution by illustrating the feasibility of measuring the cost of stockouts in a direct marketing setting. However, it also highlights two potential pitfalls when collecting these measures. The first pitfall highlights the importance of controlling for customer characteristics. The simplest approach to evaluating the cost of stockouts is to compare the behavior of customers who experienced a stockout with those who did not. Firms that make this comparison reach the puzzling conclusion that stockouts increased subsequent profits! This erroneous conclusion results from the comparison of nonequivalent groups. Customers who purchase frequently from a firm are the most likely to experience a stockout. The group of customers who experience a stockout includes many of the firm’s best customers, while the group that does not experience a stockout includes many consumers who purchase infrequently. Firms must carefully control for customer characteristics to ensure that they are comparing equivalent groups of customers.

Second, the study also highlights the importance of measuring long-run effects. Measuring long-run effects is often difficult and costly as there is, by definition, a long time between action and measurement. In our study, the design, implementation, and data collection took more than two years. Few firms have the patience or resources to implement a study and then track behavior over multiple years. Yet the findings in this study confirm that firms cannot hope to implement optimal inventory policies without measuring these long-run effects.

An online supplement to this paper is available on the Management Science website (http://mansci.pubs.informs.org/ecompanion.html).
References


