Management Science

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

Observation Bias: The Impact of Demand Censoring on Newsvendor Level and Adjustment Behavior

Nils Rudi, David Drake

MANAGEMENT SCIENCE
Vol. 60, No. 5, May 2014, pp. 1334–1345
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2013.1825

Observation Bias: The Impact of Demand Censoring on Newsvendor Level and Adjustment Behavior

Nils Rudi
INSEAD, 138676 Singapore, nils.rudi@insead.edu

David Drake
Harvard Business School, Harvard University, Boston, Massachusetts 02163, ddrake@hbs.edu

In an experimental newsvendor setting, we investigate three phenomena: *level behavior*—the decision maker's average ordering tendency; *adjustment behavior*—the tendency to adjust period-to-period order quantities; and *observation bias*—the tendency to let the degree of demand feedback influence order quantities. We find that, in three of four conditions, the portion of mismatch cost that results from adjustment behavior exceeds the portion of mismatch cost caused by level behavior. Observation bias is studied through censored demand feedback, a situation that arguably represents most newsvendor settings. When demands are uncensored, subjects tend to order below the normative quantity when they face high margin and above the normative quantity when they face low margin, but in neither case do they order beyond mean demand (the pull-to-center effect). Censoring in general leads to lower quantities, magnifying the below-normative-level behavior when they face high margin but partially counterbalancing the above-normative-level behavior when they face low margin, violating the pull-to-center effect in both cases.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2013.1825.

Keywords: inventory management; behavioral operations; decision analysis
History: Received June 3, 2009; accepted May 1, 2013, by Christian Terwiesch, operations management. Published online in *Articles in Advance* March 10, 2014.

1. Introduction and Literature Review

The newsvendor model is the fundamental model for managing inventory under demand uncertainty and has received significant research attention. Most of this research has taken a normative approach, whereas behavioral issues—the focus of this paper—have only recently received attention.

We study two aspects of ordering behavior in a repeated newsvendor model and a bias that affects them both. We define *level behavior* as a decision maker's average order quantity. Using statistical terminology, level behavior can be thought of as the first moment of ordering behavior. If, for example, the normative order quantity is 730 and a newsvendor orders 810 on average, then there is a deviation of 80 between the normative quantity and the level behavior. A prevalent level-behavior result in the literature is that, on average, subjects tend to order quantities between the expected demand and the normative order quantity; this behavior has been termed the *pull-to-center* effect. We define *adjustment behavior* as the tendency to adjust period-to-period order quantities. Again, in statistical terminology, adjustment behavior can be thought of as the second moment of ordering behavior. If, for example, one newsvendor alternates between orders of 809 and 811, and another newsvendor alternates between orders of 780 and 840, we say that the second newsvendor exhibits a greater degree of adjustment behavior than the first. Cachon and Terwiesch (2009, p. 2) give an example of such quantity adjustments with respect to influenza vaccines, for which one would expect the same ex ante season-to-season parameters and, hence, no strong reason to adjust the quantity between two consecutive seasons:

> There were 95 million doses of the flu vaccine produced for the 2002–2003 flu season in the United States. Unfortunately, 12 million doses were not used and had to be destroyed (a vaccine is good only for one flu season). Only 83 million doses of the flu vaccine were produced for the next season, 2003–2004. (Not coincidentally, $95 - 12 = 83$.) Unfortunately, in that season there were widespread shortages, leading to flu-related deaths, especially in Colorado.

This example is consistent with demand chasing, a bias that drives adjustment behavior and that has been observed frequently in the experimental newsvendor literature.

Though the literature explores both level and adjustment behavior in newsvendor settings, we are, to our knowledge, the first to estimate the performance impacts of these two behavioral patterns. We propose a method to disambiguate the overall behavioral cost (i.e., the increase of expected mismatch cost resulting from behavioral deviations from the normative order quantity) by estimating the effect of level and adjustment behavior on expected mismatch cost, which we term *level cost* and *adjustment cost*, respectively. Level and adjustment behavior are interrelated—e.g., random errors in ordering and demand chasing can lead to level effects, as shown by Su (2008) and Bostian et al. (2008), respectively—but our method serves as a useful approximation of each source of behavioral cost. Understanding the relative magnitude of these costs in different settings can help managers decide of which they should be most wary.

Schweitzer and Cachon (2000) initiated the exploration of behavioral issues in newsvendor settings, primarily focusing on level behavior. They find that order quantities tend to be greater than the normative solution in low-margin settings and less than the normative solution in high-margin settings, a pattern reported all through the experimental newsvendor literature (e.g., Bolton and Katok 2008, Lurie and Swaminathan 2009, Bostian et al. 2008, Benzion et al. 2008, Bolton et al. 2012). Schweitzer and Cachon (2000) showed that, of the 10 bias and preference structures they considered, only anchoring and insufficient adjustment and ex post inventory error minimization can explain such behavior. Since then, Bostian et al. (2008) have shown that demand chasing can lead to the pull-to-center behavior, and Su (2008) has proposed that random errors in ordering can result in such a pattern; Kremer et al. (2010) found empirical evidence refuting the latter as a sole explanation.

With respect to the treatment of adjustment behavior in the literature, Schweitzer and Cachon (2000) find weak support for demand chasing. Bolton and Katok (2008) segment subjects by (i) those whose order adjustment can be accounted for primarily by their trend in level behavior, (ii) those exhibiting a prevalence of adjustment toward the most recent demand realization (demand chasing), (iii) those tending to adjust away from the most recent demand realization (gambler's fallacy), and (iv) those whose adjustments were not statistically different from random. Other researchers have also reported demand chasing and trending in level behavior, both of which contribute to order adjustment (see, e.g., Bostian et al. 2008, Benzion et al. 2008, Lurie and Swaminathan 2009). We contribute to the understanding of adjustment behavior by providing a way to approximate its impact on performance. We find that, in three of four conditions, adjustment costs lead to significantly greater profit erosion than level costs, underscoring the importance of accounting for this form of deviation from the normative prescription.

Although various aspects of level and adjustment behavior have been reported, measures of performance applied in the literature have only considered the impacts of these behaviors in an aggregate metric. Those who report on performance often do so in terms of total profits (Benzion et al. 2008, Lurie and Swaminathan 2009), with others reporting profits resulting from level effects alone (Schweitzer and Cachon 2000) and others not reporting performance impacts at all (Feiler et al. 2013). We extend the literature by offering a way to approximate the impact of level and adjustment behavior on performance erosion. This is useful in (i) developing the understanding that both behaviors drive costs and (ii) assessing the relative magnitudes of these costs in various settings.

Arguably, in most practical settings, demand feedback tends to be in the form of demand realizations censored by the ordered quantity; i.e., newsvendors observe sales realizations, rather than demand realizations. Furthermore, even when firms have data that could be used to develop a rough estimate of lost sales, they do not, in our experience, tend to do so. The earlier example of the influenza vaccine reports a specific leftover quantity for the 2002–2003 season, but only concludes that the shortages in the 2003–2004 season were "widespread." We define *observation bias* as the tendency of the degree of feedback available—here, whether full demand realizations are observed or demands are censored—to influence level and adjustment behavior. We manipulate demand feedback (censored and uncensored) as an experimental treatment to induce and test for observation bias. We find that demand censoring induces a reduction in order quantities relative to uncensored settings, which provides support for an observation-bias effect on level behavior. We also investigate the impact of observation bias on adjustment behavior, finding that censoring attenuates adjustment; i.e., the magnitude of adjustments are related to the degree of variability observed by subjects. By providing a method to tie level and adjustment effects to performance (in terms of estimated mismatch costs), these results can help managers prioritize order-quantity improvements based on product margins and the degree of demand feedback available in the setting in which they operate.

In normative models prescribing ordering policies in repeated settings when the underlying demand distribution is unknown (in contrast to the setting in our study) and demands are censored, it is common for censoring to lead to higher normative order levels because of the value of improved demand information through the ability to observe a larger range of demand realizations (Harpaz et al. 1982, Nahmias 1994, Lariviere and Porteus 1999, Ding et al. 2002). Feiler et al. (2013) study judgment in censored environments. Of their four studies, the fourth considers sequential risk taking in a dynamic environment and is hence not directly relevant to our study. Their studies 1 and 2 are newsvendor settings where the censoring level (order quantity) is exogenously given and subjects estimate the (unknown) mean of random demand. They manipulate censoring level and whether the censoring level is fixed or variable (study 1) and the underlying demand variability (study 2). Feiler et al. (2013) find that mean estimates tend to be between the true mean (which they demonstrate can be closely approximated by the likelihood based approach of Nahmias 1994) and the average of censored observations. Their study 3 considers the dual tasks of mean demand estimation and newsvendor decision making. However, their unit underage and unit overage costs are identical, reducing the newsvendor decision problem in the absence of learning to that of estimating mean demand. They manipulate mean demand knowledge (whether subjects are informed about the true mean) and demands censoring. In the conditions where the true mean is known, subject estimates of mean demand are, not surprisingly, indistinguishable from the true mean. Similarly, with an unknown mean and uncensored demand, subject estimates of mean demand were not significantly different from the true mean. However, in the condition with an unknown mean and demand censoring, subject estimates of mean demand are significantly less than the true mean but more than observed average censored demands. In all four cases, mean demand estimates and stocking decisions are not significantly different. Therefore, Feiler et al. (2013) detect bias in the mean estimation task when mean demand is unknown and censored, but it is arguable whether they detect bias in the quantity decision task in any of the four settings. Ordering the estimated mean is optimal when the demand distribution is known or when the demand distribution is unknown and there is no demand censoring. When the demand distribution is unknown and demands *are* censored, the normative quantity is greater than the estimated mean demand in early periods and converges toward the estimated mean demand with each successive period. Feiler et al. (2013) do not explicitly test order decisions against these normative quantities, but it seems likely that there would be no significant bias detected in the quantity decision (relative to a normative baseline) based on the results that they do report.

In contrast to Feiler et al. (2013), we focus on the impact of demand censoring on newsvendors' quantity decisions when overage and underage costs are unequal. This setting reflects the fundamental newsvendor trade-off, balancing the risk of having too little inventory versus the risk of having too much, when the optimal order quantity differs from mean demand. With a known demand distribution, we detect significant bias in the quantity decision task as a result of both the pull-to-center effect and demand censoring. This contrasts Feiler et al. (2013), where order quantities with a known demand distribution, equal overage, and underage costs do not differ significantly from the normative quantity (i.e., mean demand). As discussed, we also provide a way to estimate the expected cost of both level and adjustment behavior, and we explore these behavioral costs and total mismatch costs in both censored and uncensored settings.

Section 2 outlines the normative newsvendor model. Section 3 describes the paper's experimental design and the notation for each experimental condition. In §4, we present the theory and hypotheses (§4.1) and results (§4.2) corresponding to level and adjustment behavior and observation bias. Section 4.3 presents additional findings related to learning in terms of both level and adjustment behavior and the relationship between both behaviors. Section 5 provides a summary discussion of key results and managerial implications.

2. The Normative Newsvendor

In the normative model, the newsvendor decides order quantity $\mathbf{Q}$ at unit cost w while facing uncertain demand D in a single-period environment. After demand is realized, the newsvendor sells, at unit revenue r, a quantity limited by demand D and order quantity $\mathbf{Q}$. Here we consider the situation in which leftover inventory is disposed of (at no value). Define $a^+ = \max(0, a)$. The newsvendor's decision then results in realized profit

$$\Pi(D, \mathbf{Q}) = (r - w)D - \mathbf{G}(D, \mathbf{Q}), \tag{1}$$

where $\mathbf{G}(D, \mathbf{Q}) = c_u(D - \mathbf{Q})^+ + c_o(\mathbf{Q} - D)^+$ is the mismatch cost when ordering $\mathbf{Q}$, with parameters unit underage cost $c_u = r - w$ and unit overage cost $c_o = w$. Hence, the profit function (1) is separated into two parts: (i) profit if one could obtain exact demand information before committing to the quantity and (ii) the cost of demand uncertainty (mismatch cost). The latter consists of (a) the financial consequence per unit of unmet demand (unit underage cost) times the expected unmet demand and (b) the financial consequence per unit of leftover quantity (unit overage cost) times the expected leftover quantity. Let $\Pi(\mathbf{Q}) = \mathbb{E}\,\Pi(D, \mathbf{Q})$ denote expected profit and $\mathbf{G}(\mathbf{Q}) = \mathbb{E}\,\mathbf{G}(D, \mathbf{Q})$ denote expected mismatch cost when ordering $\mathbf{Q}$. Note that expected mismatch cost is the *controllable* part of expected profit, i.e., the only part that depends on the decision variable $\mathbf{Q}$. Hence, the minimizer of $\mathbf{G}(\mathbf{Q})$ will be equal to the maximizer of $\Pi(\mathbf{Q})$. The normative solution minimizing expected mismatch cost when demand follows a continuous distribution is the order quantity $\mathbf{Q}^*$ that satisfies $F(\mathbf{Q}) = c_u/(c_u + c_o)$, where the left-hand side $F(\cdot)$ is the cumulative distribution function of D and the right-hand side is the newsvendor fractile.

3. Experimental Design and Notation

We conducted newsvendor experiments as one session of the core logistics course in the undergraduate program at the Norwegian School of Management. Each session started with a 90-minute lecture on the newsvendor model that included an explanation of the key features of the problem, the motivation of its relevance, an explanation of the trade-off between underage and overage (including which side of the mean to order on), and insight into the effects of parameters. The lecture did not, however, include the newsvendor fractile or a method for computing the normative quantity. The lecture also specifically covered the issue of censored demand realizations and subjects were informed that some—but not all—of them would face such situations and that all of them would be informed about the actual underlying demand distribution that was independent and identically distributed (i.i.d.) across periods. Similarly, they were informed that some would face conditions in which margin exceeded cost, and others would face conditions in which cost exceeded margin. After the lecture, students participated in a 2×2 between-subject experiment with treatments to manipulate the margin of "widgets" sold (low versus high) and the degree of demand feedback (uncensored versus censored).

In the experiment, we used normally distributed demands with mean $\mu = 1{,}000$ and standard deviation $\sigma = 400$, so the coefficient of variation was large enough to make an impact and small enough for a normal distribution to be reasonable. All subjects experienced realizations from a common demand path. Since all subjects also received complete information regarding the demand distribution up front, we can focus exclusively on the impact of observation bias, rather than confounding its effect with that of demand learning. Furthermore, a known demand distribution makes a myopic policy optimal, thus avoiding complications that would result from dynamics.[1]

We fixed unit revenue at $r = 12$ and manipulated unit cost as one treatment, with $w = 3$, which we denote *high-margin good* (HMG), and $w = 9$, which we denote *low-margin good* (LMG). These parameters result in an HMG critical fractile of 0.75 and an LMG critical fractile of 0.25, which aligns with several studies (e.g., Schweitzer and Cachon 2000, Bolton and Katok 2008, Bostian et al. 2008, Lurie and Swaminathan 2009), but, notably, differs from Feiler et al. (2013), who use a fractile of 0.5. We study LMG and HMG settings—i.e., fractiles below and above 0.5—for two main reasons. First, the fundamental newsvendor trade-off is between ordering too little and ordering too much with, in general, asymmetric consequences. Hence, we contribute to Feiler et al. (2013), where only equal underage and overage costs are considered (comparisons are provided following the development of Hypothesis 2 and its related results in §4). Second, we expect that censoring will affect level and adjustment behavior and costs differently in LMG and HMG settings, which we detail in our hypotheses.

To study observation bias, we manipulated demand feedback with *uncensored* demand, where subjects observed actual demand realizations D, and *censored* demand, where subjects were limited to observing quantities sold, $\min(D, \mathbf{Q})$. Subjects in the uncensored condition were also informed of underage costs, $c_u(D - \mathbf{Q})^+$, and overage costs, $c_o(\mathbf{Q} - D)^+$, for each period. In the censored condition, subjects were still informed of demand outcomes and overage costs for periods following overage outcomes, but they were only informed that a stockout occurred during periods in which they experienced underage.

A total of 310 subjects participated in the experiment, each randomly assigned to one of the four conditions. We include the results from 269 of these subjects in the results presented in §4. Of the other 41 subjects, 30 were dropped from the study because they did not complete the 50 trials in the allotted time, 2 were unable to play the game because of technical problems, 4 were excluded because of strong evidence suggesting that they had obtained demand information, and 5 were excluded for severe overordering. (Of those five, four had orders in excess of 5,000 and one had an average order quantity 8.5 standard deviations greater than the condition's average.) We also corrected order quantities in three instances in which a mistyping was obvious. Subjects completing all 50 trials entered a lottery in which each "euro" of profit they accumulated throughout the experiment earned them a chance to win an iPod. One iPod was raffled in each of the six course sections.

We use expected mismatch cost (cost of demand uncertainty) rather than profit to estimate subject performance. We do this for two reasons. First, it represents the *controllable part* of expected profit—the part that depends on $\mathbf{Q}$. Second, it enables us to compare the alternatives with LMGs and HMGs on equal terms (since the expected mismatch cost functions of the two margin scenarios are mirror images of each other around mean demand). In the general model—and when making statements that are valid independent of margin—we use capital and boldface notation for quantity decisions $\mathbf{Q}$ and mismatch cost $\mathbf{G}$. When making margin-specific statements, we use regular uppercase notation for HMG and lowercase notation for LMG (a notation mnemonic: big notation corresponds to big margin and small notation corresponds

[1] For robustness, we also ran an experiment, described in §4.3, with a known and an unknown demand distribution.

to small margin). The normative order quantities are then $q^* = 730$ and $Q^* = 1{,}270$ for LMG and HMG, respectively. Furthermore, we use superscript u for the uncensored case and superscript c for the censored case. We index a specific subject (newsvendor) by i and consider the situation in which there are $T = 50$ repeated periods, indexed by t, and D_t's are normal i.i.d. Dropping the subscript t indicates an average value over t. So subject i's average quantity when facing HMG is $Q_i = \sum_{t=1}^{T} Q_{i,t}/T$, subject i's standard deviation of order quantities when facing LMG is σ_{q_i}, and subject i's average expected mismatch cost when facing LMG is $g_i = \sum_{t=1}^{T} g(q_{i,t})/T$.

Based on the experiment's 2×2 design, each subject is a member of one of four disjoint sets determined by condition—LU, LC, HU, and HC—where L denotes a low-margin good, H denotes a high-margin good, U denotes uncensored demand, and C denotes censored demand. (Each subject plays one of these conditions throughout the experiment.) We denote condition average measures by dropping the subscript i; e.g., $Q^c = \sum_{i \in HC} Q_i / \langle HC \rangle$ and $\sigma_{q^u} = \sum_{i \in LU} \sigma_{q_i} / \langle LU \rangle$, where $\langle A \rangle$ denotes the cardinality of the set A (i.e., the number of subjects in the condition).

4. Development and Analysis of Hypotheses

In §4.1, we develop theory and hypotheses related to level and adjustment behavior with and without censoring and their expected effects on associated costs. We present the corresponding results in §4.2 and additional results and robustness tests in §4.3.

It is important to note the unit of analysis on which we generally measure and test ordering behavior in this paper. When performing tests on level behavior (average order quantities), we do so on the population of subjects' average quantities $\mathbf{Q}_i$. The rationale behind this is that it is individual subjects, not groups of subjects, who exhibit a particular behavioral pattern. Similarly, when performing tests on the effect of adjustment behavior on decisions (variability of order quantities), we do so on the population of within-subject standard deviations of quantities $\sigma_{\mathbf{Q}_i}$ since within-subject variability represents the extent to which a subject adjusts her order quantities, whereas between-subject variability reflects the extent to which subjects differ in level behavior.

4.1. Theory and Hypotheses

4.1.1. Level Behavior in Uncensored Environments. Here we study subject *decisions* (i.e., order quantities) with respect to level behavior when demands are uncensored.

Schweitzer and Cachon (2000) find support for the pull-to-center effect, i.e., a tendency to order between the normative solution and the expected demand, which is consistent with anchoring and insufficient adjustment as well as with ex post inventory error minimization. We expect that our experiment will yield corresponding results when demand is uncensored.

Hypothesis 1 (H1). *The order quantities with uncensored demands will fall between the normative quantity and mean demands,* (a) $q^* < q^u < \mu$ *and* (b) $\mu < Q^u < Q^*$.

4.1.2. The Effect of Observation Bias on Level Behavior. Most practical newsvendor settings involve demand censoring. Firms typically have visibility to sales but rarely have the luxury of observing true demand realizations. We posit that censoring will impact observation bias, resulting in reduced order quantities relative to those observed when demand is uncensored.

We propose that demand censoring can induce observation bias through salience differences that arise from the asymmetric awareness of overage and underage costs. Lewis (1969, p. 35) defines salience as an entity or property that "stands out from the rest by its uniqueness in some conspicuous respect." Under demand censoring, decision makers receive more precise demand feedback following overage events than following underage events; i.e., they receive information that is asymmetric in its salience. We posit that this asymmetry can affect decision making through the salience effect, which Tversky and Kahneman (1974) offer as one factor that can influence subjective probabilities under the availability heuristic. Tversky and Kahneman (1973, p. 207) define the availability heuristic as "a judgmental heuristic in which a person evaluates the frequency of classes or the probability of events by availability, i.e., by the ease with which relevant instances come to mind."

Through the availability heuristic, decision makers place greater subjective probabilities on events or outcomes that are more easily recalled. It is important to note that Tversky and Kahneman (1973) point out that the ease with which relevant information *could* be recalled from memory affects subjective probabilities; the cognitive operations of retrieval (recall from memory) and construction (solving the task at hand) are not necessary for the heuristic to affect decisions. Salience-effect bias arises when, because of asymmetries in the richness or cognitive impact of informational cues, the availability heuristic leads to subjective probability distortion relative to objective probabilities. Taylor et al. (1979, p. 367) once argued that salience effects impact decision making on the subconscious level, i.e., that they are "an automatic perceptual bias (not unlike optical illusions), ... which occur without intention." A softer stance is taken in Fiske and Taylor (1991), who state that salience effects are not fully automatic because they can be overcome through training, coaching, and other "forms of involvement." This implies that subjects do not have to explicitly consider the subjective probabilities that they place on various demand realizations for the salience effect to impact their decision making. Simply being exposed to the asymmetric information cues is sufficient.

In light of the discussion on observation bias, we expect demand censoring to affect level behavior according to the following hypotheses:

Hypothesis 2 (H2). *Demand censoring will lead to a reduction in order quantity levels*, (a) $q^c < q^u$ *and* (b) $Q^c < Q^u$. *The difference caused by demand censoring will be more prominent for LMG than for HMG*, (c) $Q^u - Q^c < q^u - q^c$.

Hypothesis 2(c) arises from the conjecture that order quantities will be greater for HMG than for LMG. This level difference implies that demand feedback will be censored more frequently for LMG than for HMG. Hence, we expect a stronger censoring effect for LMG than for HMG.

Should these hypotheses hold, demand censoring will cause subjects to order further from the normative quantity in HMG settings. In LMG settings, however, demand censoring would cause subjects to order quantities toward (and possibly beyond) the normative quantity q^*; i.e., it would partially counterbalance the pull-to-center effect on q^u.

As discussed in §1, Feiler et al. (2013) consider in their study 3 a newsvendor trade-off of equal underage and overage costs. In such a setting, the normative quantity is—in the absence of learning—equal to the mean demand estimate. Feiler et al. (2013) posit that censoring will not have a significant impact on order quantities when mean demand is known (i.e., order quantities will equal the true mean when this is provided to subjects). In contrast, we posit in H2 that, with a nontrivial newsvendor trade-off (nonequal underage and overage costs), censoring will lead to significantly lower order quantities.

4.1.3. The Effect of Observation Bias on Adjustment Behavior. It is reasonable to expect that subjects' quantity adjustments will be influenced by what they observe from period to period, which, in turn, depends on demand realizations. When demand feedback is uncensored, subjects can observe full demand realizations, D. When demand feedback is censored, however, subjects' period-to-period observations are restricted to sales realizations, $\min(D, \mathbf{Q})$. The standard deviation of demand in our setting is $\sigma = 400$, whereas the standard deviation of sales is increasing in $\mathbf{Q}$ and is bounded above by 400.

As was the case with level behavior, subjects should normatively ignore sales and demand signals, given full knowledge of the system parameters. However, as Massey and Wu (2005) illustrated through their system neglect hypothesis, subjects are likely to place inordinate weight on signals. Therefore, we conjecture that the more variability there is in what subjects observe, the more they will adjust their order quantities. We therefore expect that demand censoring will lead to a lower degree of adjustment of order quantities and that this effect will be stronger for LC than for HC, since we expect the former to have lower order quantities and therefore less variable sales. Note that this relationship should be interpreted with care because of its natural endogeneity; i.e., we conjecture that the variability of order quantities depends on the variability of sales, but the variability of sales also depends on the variability of order quantities.

Hypothesis 3 (H3). *Orders will be less variable with censored demand feedback than with uncensored demand feedback*, (a) $\sigma_{q^c} < \sigma_{q^u}$ *and* (b) $\sigma_{Q^c} < \sigma_{Q^u}$. *The effect of demand censoring on order variability will be more prominent for LMG than for HMG*, (c) $\sigma_{q^u} - \sigma_{q^c} > \sigma_{Q^u} - \sigma_{Q^c}$.

4.1.4. Performance Impacts. To investigate the *effect* of level and adjustment behavior with and without demand censoring, we consider their estimated impacts on expected mismatch cost. In this respect, the expected additional cost due to subject i's level behavior is denoted $\mathbf{\Delta}_i$ and is estimated by the difference between the expected mismatch cost when ordering the subject's average quantity and the expected mismatch cost when ordering the normative quantity, $\mathbf{\Delta}_i = \mathbf{G}(\mathbf{Q}_i) - \mathbf{G}(\mathbf{Q}^*)$.[2] Further, the expected cost of subject i's adjustment behavior is denoted $\mathbf{\Psi}_i$ and is estimated by the difference between the subject's average expected mismatch cost and the expected mismatch cost when ordering the average order quantity, $\mathbf{\Psi}_i = \mathbf{G}_i - \mathbf{G}(\mathbf{Q}_i)$. It follows trivially that $\mathbf{\Delta}_i \geq 0$, with equality iff $\mathbf{Q}_i = \mathbf{Q}^*$. By Jensen's inequality, it follows that $\mathbf{\Psi}_i \geq 0$, with equality iff $\mathbf{Q}_{i,t} = \mathbf{Q}_i, \forall t$. To summarize, subject i's average expected mismatch cost is disambiguated into the normative cost and the estimates of the level cost and the adjustment cost, $\mathbf{G}_i = \mathbf{G}(\mathbf{Q}^*) + \mathbf{\Delta}_i + \mathbf{\Psi}_i$. We refer to $\mathbf{\Delta}_i + \mathbf{\Psi}_i$ as the *total behavioral cost*.

For HMG, based on H1(b) and H2(b), we expect that the level cost will be higher in censored than in uncensored demand feedback conditions. This effect is further magnified in terms of the effect on expected mismatch cost, since, by convexity of $\mathbf{G}$, the further $\mathbf{Q}$ is from the normative order quantity $\mathbf{Q}^*$, the steeper $\mathbf{G}(\mathbf{Q})$ becomes. For LMG, based on H1(a)

[2] Note that we continue to follow the convention of using uppercase and boldface notation when not being margin specific, nonbold uppercase for HMG, and nonbold lowercase for LMG.

and H2(a), we conjecture that demand censoring will partially counterbalance the pull-to-center effect but may reduce orders to levels below the normative solution. Consider, then, the case in which demand censoring leads to order quantities below the normative quantity; i.e., $q_i^c < q^*$. By the fact that, for LMG,

$$\frac{d}{dx}[g(q^*+x)-g(q^*-x)] = (c_u+c_o)[F(q^*-x)+F(q^*+x)-2F(q^*)]>0, \quad \forall x>0,$$

it follows that $g(q^*-x) < g(q^*+x)$, $\forall x > 0$; i.e., the mismatch cost is "flatter" to the left of q^* than to the right. Therefore, if demand censoring reduced order quantities beyond the normative solution for LMG (i.e., if $q^c < q^*$), then δ^c would be less than δ^u as long as the demand-censoring effect were not so severe as to reduce quantities to a level more distant from q^* than they were without demand censoring. On the other hand, if demand censoring reduces order level as expected, but not to levels below q^*, then $\delta^c < \delta^u$ would hold trivially. Hence, it is reasonable to assume that censoring will reduce level costs in LMG settings.

The effect on adjustment cost will predominantly depend on the degree of order adjustment (i.e., within-subject variability of order quantities) and the degree of convexity of the expected mismatch cost in the region of the order quantities (more convexity will lead to a stronger effect on adjustment cost). The degree of convexity of the expected mismatch cost is quantified by its second derivative, which takes its maximum value at a quantity equal to mean demand μ and is decreasing symmetrically on both sides of μ. In both margin treatments considered here, $G''(Q)$ reduces by about 20% from its maximum at the pull-to-center boundary $\mathbf{Q} = \mu$ to the other pull-to-center boundary $\mathbf{Q} = \mathbf{Q}^*$.

When comparing adjustment cost between the censored and uncensored demand feedback conditions, we conjecture that the effect of the difference in order variability will dominate the relatively small effect of the degree of convexity, as discussed. For LMG, we expect from H2(a) that the mismatch cost will be less convex in the region of q^c than in the region of q^u. Furthermore, we expect from H3(a) that subjects' order quantities when demand feedback is censored, q^c, will be less variable than their quantities when demand feedback is uncensored, q^u. Both of these effects suggest that the adjustment cost will be lower under censored demand feedback than under uncensored demand feedback. For HMG, in contrast, we expect from H1(b) and H2(b) that the effect of G's degree of convexity will contribute to there being a larger adjustment cost when demands are censored than when demands are uncensored. However, H3(b) suggests that the effect of quantity variability will be in the opposite direction. From this discussion, we expect that, for HC, the latter effect will dominate the former and conjecture that the adjustment cost will be lower in HC than in HU.

Since the effect of demand censoring is expected to be in the same direction for both LMG behavioral costs, we clearly expect the overall behavioral cost to be smaller in LC than in LU. When comparing HC and HU, however, the effect of demand censoring on the two behavioral costs is expected to be in different directions, and it is unclear which one will dominate. Hence, we cannot make a hypothesis regarding the direction of overall behavioral cost.

Hypothesis 4 (H4). *Demand censoring will lead to a smaller level cost for LMG and a larger level cost for HMG,* (a) $\delta^u > \delta^c$ *and* (b) $\Delta^u < \Delta^c$. *Demand censoring will lead to lower adjustment costs,* (c) $\psi^u > \psi^c$ *and* (d) $\Psi^u > \Psi^c$. *Demand censoring will lead to a smaller total behavioral cost for LMG,* (e) $\delta^u + \psi^u > \delta^c + \psi^c$.

4.2. Results

The descriptive statistics resulting from the experiment are given in Table 1.

4.2.1. Results for Level Behavior in Uncensored Environments. The average quantity values are shown in Figure 1(a).

To test H1(a) and H1(b), we construct 99% confidence intervals around q^u and Q^u. The null hypotheses (i.e., that pull-to-center will not be observed) will be rejected at the $p < 0.01$ level if the confidence limits are contained within their pull-to-center regions, which have bounds established by expected demand and the normative order quantity (i.e., between 730 and 1,000 in LMG and between 1,000 and 1,270 in HMG). From Table 2, we see that there is statistically significant evidence for subjects' average order quantities falling within the pull-to-center boundaries ($p < 0.01$ for both LMG and HMG). These results are consistent with Schweitzer and Cachon (2000), as well as with the newsvendor experiments referenced in our literature review.[3]

4.2.2. Results for the Effect of Observation Bias on Level Behavior. The average quantity values of the four conditions are given in Table 1 and shown in Figure 1(a).

[3] Schweitzer and Cachon (2000) also model other preference functions with a clear prediction of effect on order quantities. Specifically, risk seeking and stockout averse preferences would lead to overordering, whereas risk aversion, underestimation of mismatch cost, waste aversion, and loss-aversion preferences would lead to underordering. The latter four preferences as secondary effects would lead us to conjecture that the pull-to-center effect is weaker for LMG than for HMG, which in notation would be $(Q^* - Q^u) - (q^u - q^*) > 0$. This is supported here ($p < 0.01$) and is in contrast to Schweitzer and Cachon (2000), although in their case not with significance.

Table 1 Descriptive Statistics

	LU	LC	HU	HC
Conditions	$I = 64$	$I = 68$	$I = 66$	$I = 71$
		Subjects' average order quantities		
Normative: Q^*	730	730	1,270	1,270
Mean	836.72	690.31	1,080.76	964.13
Median	844.05	687.70	1,053.63	1,043.54
Std. deviation	118.04	146.66	140.69	236.70
95% CI	[807.24, 866.21]	[654.81, 725.81]	[1,046.17, 1,115.34]	[908.11, 1,020.16]
Min	526.26	387.00	808.84	515.90
Max	1,156.98	1,089.48	1,466.20	1,445.00
		Within-subject standard deviation		
Mean	245.67	176.13	270.54	211.65
Median	246.78	157.97	271.74	195.92
Std. deviation	82.33	81.15	85.10	90.73
95% CI	[225.10, 266.23]	[156.49, 195.77]	[249.62, 291.46]	[190.17, 233.13]
Min	40.28	56.67	71.17	4.24
Max	445.83	479.11	462.39	519.42
		Expected mismatch cost		
Normative: $G(Q^*)$	1,525.33	1,525.33	1,525.33	1,525.33
Mean	1,993.34	1,800.87	2,230.64	2,602.65
Median	1,875.58	1,728.86	2,260.59	2,336.81
Std. deviation	341.65	262.26	404.70	889.46
95% CI	[1,907.99, 2,078.68]	[1,737.39, 1,864.35]	[2,131.15, 2,330.12]	[2,392.12, 2,813.18]
Min	1,542.19	1,547.75	1,595.08	1,525.41
Max	3,489.14	3,101.20	3,097.31	4,665.81

Note. CI, confidence interval.

From Table 3, we see that for both LMG and HMG order quantities tend to be significantly lower with demand censoring than without it (by 141.41 and 116.62 units, respectively), providing support for H2(a) and H2(b) ($p < 0.01$ for both LMG and HMG). In fact, we find the effect of demand censoring to be so strong that subjects' ordering violates the lower bound of the pull-to-center regions—a robust result in the literature—for both LC (two tailed t-test, $p < 0.05$) and HMG, although not significantly so in the latter case. We do not find significance for H2(c), but it holds directionally. (Note that, unlike the other tests here, the test of H2(c), a difference-in-difference test, is performed using ANOVA with contrasts.)

Table 2 99% Confidence Interval of Uncensored Quantities Testing Pull-to-Center Effect

	Mean	SE	99% CI
H1(a) LU	836.72	14.76	[797.53, 875.91]
H1(b) HU	1,080.76	17.32	[1,034.80, 1,126.71]

Note. CI, confidence interval.

These results contrast the finding in Feiler et al. (2013) with equal underage andoverage costs and

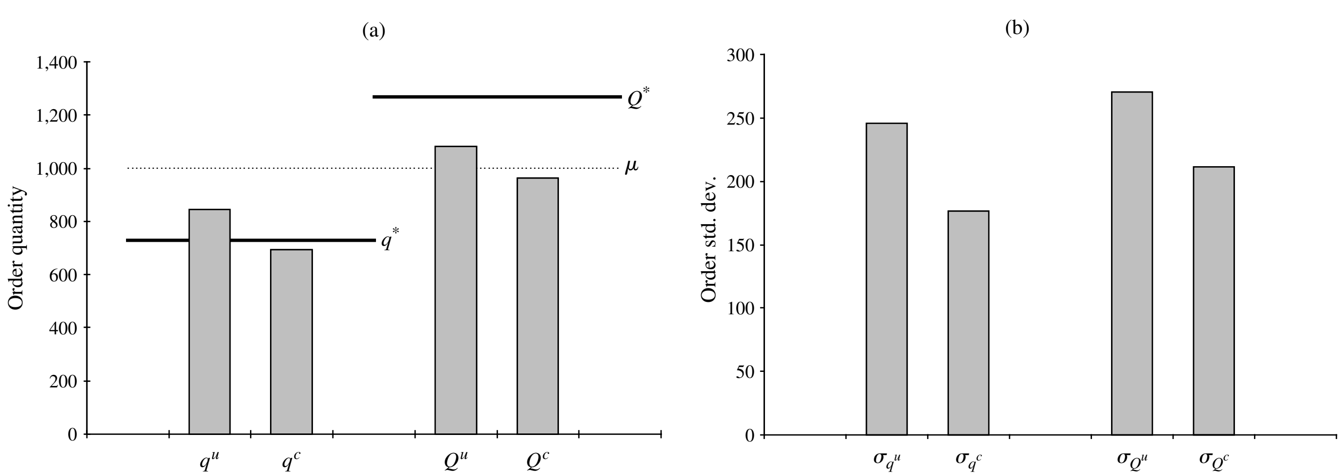


Figure 1 (a) Average Order Quantities and (b) Average Standard Deviation of Order Quantities for the Four Conditions

Table 3 *t*-Tests of the Effect of Censoring on Level Behavior

	Uncensored	Censored	Difference	
H2(a) LMG q	836.72	690.31	146.41	$t = 6.34$, p-value $= 0.0000$
H2(b) HMG Q	1,080.76	964.13	116.62	$t = 3.53$, p-value $= 0.0003$
H2(c) Difference in differences			29.79	$t = 0.72$, p-value $= 0.2348$

given mean, where censoring did not significantly influence the order quantities. This suggest that when the newsvendor trade-off is nontrivial, i.e., when underage and overage costs are nonequal, subjects become biased from the censoring and that the result of Feiler et al. (2013) is contingent on equal underage and overage costs.

4.2.3. Results for the Effect of Observation Bias on Adjustment Behavior. As discussed previously, we focus on within-subject variability to assess the magnitude of adjustment behavior. We do so because subjects rather than groups are the unit of analysis.

The average within-subject quantity standard deviations of the four conditions are shown in Table 1 and displayed in Figure 1(b).

From Table 4, we see that demand censoring leads to significantly less order variability for both LMG and HMG (by 69.54 and 58.88 units, respectively), providing support for H3(a) and H3(b) ($p < 0.01$ for both cases). To test H3(c), we used ANOVA with contrasts. Although the results directionally support the hypothesis that demand censoring has a stronger attenuating effect on order variability in LC than in HC, that support is not statistically significant.

Although the data support H3(a) and H3(b), it is unclear whether or not demand censoring tends to attenuate order adjustment as a result of sales variability being less than demand variability (i.e., as posited) or for some other reason. As an alternative explanation, order adjustment could tend to increase in order level. By H2(a) and H2(b), there was strong support for the conjecture that demand censoring tends to reduce order levels. Therefore, if order adjustment does tend to increase in order level, demand censoring could have an indirect attenuating effect on order adjustment; i.e., demand censoring tends to reduce order levels, which, in turn, could tend to reduce order adjustment.

Table 4 *t*-Tests of the Effect of Censoring on Adjustment Behavior

	Uncensored	Censored	Difference	
H3(a) LMG σ_q	245.67	176.12	69.54	$t = 4.88$, p-value $= 0.0000$
H3(b) HMG σ_Q	270.54	211.65	58.88	$t = 3.92$, p-value $= 0.0001$
H3(c) Difference in differences			10.65	$t = 0.51$, p-value $= 0.3041$

To test for such a level effect as an alternative, we focus on the uncensored demand feedback conditions. Subjects in both HU and LU observe demand realizations, so both are exposed to full demand variability (as opposed to sales variability). Sales variability would therefore not explain the observed differences in order adjustment between HU and LU. However, HU and LU do differ in average order level. Therefore, we isolate a potential level effect on order adjustment by comparing these two conditions. We find that order adjustment is significantly less in LU than in HU (from Table 1, the within-subject standard deviation difference is 24.87, which is significant at $p < 0.05$).[4] This indicates support for a level effect on adjustment behavior: A subject's order adjustment tends to increase in her average order level. This in turn provides support for an indirect effect of demand censoring on order adjustment: demand censoring tends to reduce average level, which, in turn, tends to attenuate order adjustment.

The presence of this indirect effect of demand censoring on order variability does not, however, preclude the existence of the posited direct effect as well. To test for the possibility that demand censoring has both a direct and an indirect effect on order variability, we estimated the following model through ordinary least squares:

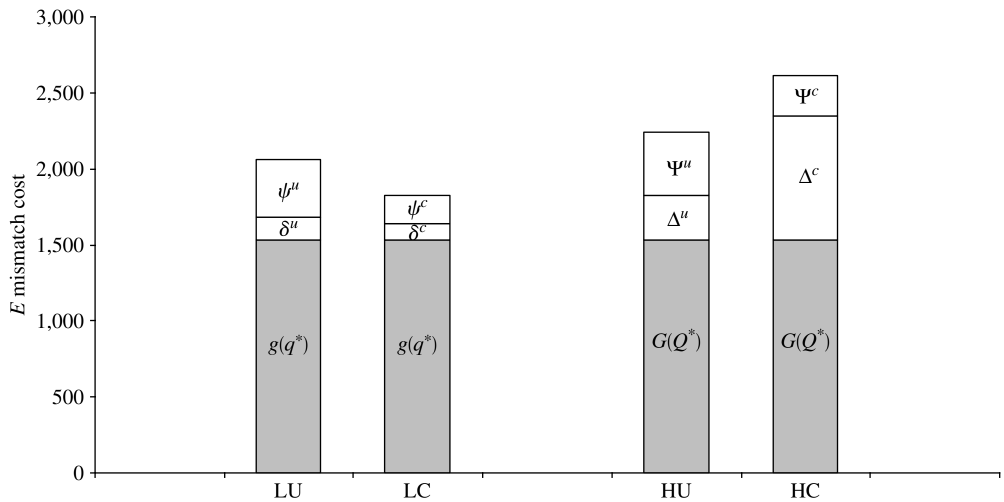
$$\sigma_{\mathbf{Q}_i} = \beta_0 + \beta_1 \mathrm{Cen}_i + \beta_2 \mathrm{HMG}_i + \beta_3 \mathrm{Cen}_i \mathrm{HMG}_i + \beta_4 \mathbf{Q}_i + \varepsilon_i,$$

where Cen_i is an indicator equal to 1 if subject i participated in LC or HC, and HMG_i is a control indicator equal to 1 if subject i participated in HU or HC. We found strong significance for Cen_i ($\beta_1 = -51.94$, $p < 0.01$) and for Q_i ($\beta_4 = 0.120$, $p < 0.01$), while HMG_i and $\mathrm{Cen}_i\mathrm{HMG}_i$ were insignificant. This suggests that demand censoring has both a direct attenuation effect on order variability and an indirect effect through order level.

4.2.4. Results for Performance Impacts. Figure 2 shows the expected normative mismatch cost, average level cost, and average adjustment cost for all four conditions.

[4] We do not include a corresponding analysis comparing HC to LC because the average observed variability (sales variability) is less in LC than in HC (a result of average order quantity being less in LC than in HC). Therefore, we cannot determine if the decreased order adjustment in LC relative to HC (a difference of 35.52, significant at the $p < 0.01$ level) is a result of reduced signal variability or reduced average level. Since demand variability is independent of order quantity, the comparison between HU and LU is not confounded in this manner.

Figure 2 Separation of Behavioral Costs for the Four Conditions



As indicated in Table 5, we find directional support for H4(a); demand censoring leads to lower level cost for LMG (by 27.21), but not with significance. For HMG, however, demand censoring leads to significantly larger level cost (by 517.95), supporting H4(b) ($p < 0.01$). Supporting H4(c) and H4(d), adjustment costs are significantly smaller under demand censoring for both LMG and HMG, by 165.26 and 145.94, respectively ($p < 0.01$ for both LMG and HMG). Interestingly, in three of the four conditions—namely LU, LC, and HU—we find that adjustment cost exceeds the corresponding level cost significantly (by 205.30, 67.30, and 113.46, respectively, with two-tailed t-test, $p < 0.01$ for LU and $p < 0.05$ for LC and HU). For HC, we find the opposite to be true; level cost exceeds adjustment cost with significance (by 550.44, with two-tailed t-test, $p < 0.01$). This highlights the importance of explicitly considering adjustment behavior in newsvendor settings in addition to the more commonly emphasized level behavior.

Finally, supporting H4(e), demand censoring leads to a lower behavioral cost by 191.50 for LMG ($p < 0.01$). Although competing level and adjustment effects prevent us from formulating a hypothesis regarding HMG total behavioral cost, we found, using a two-sided test, that demand censoring leads to a higher behavioral cost by 372.01 for HMG ($p < 0.01$).

4.3. Additional Results and Robustness

Several additional results and robustness tests were performed as a part of this research.

In terms of adjustment behavior, we find—in line with existing literature—that subjects tend to chase demand. We also find that in the censored demand environments (in which subjects experience different observed variability in terms of sales), higher observed variability tends to lead to higher adjustment behavior. Furthermore, there tends to be a positive association between level behavior and adjustment behavior, indicating that "those who get it, really get it." We also study learning in terms of level behavior (tending to get closer to normative quantity) and in terms of adjustment behavior (tending to reduce order variability). In terms of level behavior, for low-margin products this only holds directionally, but for high-margin products, it holds with significance.

On a referee's request, we also studied data from a second experiment that included an unknown and a known demand distribution as an additional treatment, resulting in a $2 \times 2 \times 2$ study design. As discussed in the introduction, normative theory gives a clear prescription that one should overorder relative to the myopic solution when facing an unknown demand distribution and censored demand.

Table 5 *t*-Tests for Behavioral Costs

		Uncensored	Censored	Difference		
H4(a)	LMG δ	131.34	104.12	27.21	$t = 0.95$,	p-value $= 0.1704$
H4(b)	HMG Δ	295.93	813.88	−517.95	$t = -4.22$,	p-value $= 0.0001$
H4(c)	LMG ψ	336.67	171.42	165.26	$t = 4.76$,	p-value $= 0.0000$
H4(d)	HMG Ψ	409.38	263.44	145.94	$t = 3.72$,	p-value $= 0.0001$
H4(e)	LMG $\delta + \psi$	468.01	275.54	192.47	$t = 3.64$,	p-value $= 0.0002$
	HMG $\Delta + \Psi$	705.31	1,077.32	−372.01	$t = -3.11$,	p-value $= 0.0023$*

*Two-sided test.

The results presented throughout this section were highly robust when testing the same hypotheses in the unknown demand settings. Furthermore, when comparing subjects' level and adjustment behaviors, as well as the components of behavioral costs, no results were statistically different between the known and unknown demand distribution in the new experiment. This suggests that our findings are robust to unknown demand distribution settings. Subjects faced with an unknown demand distribution and censored demands tend to behave contrary to what normative theory suggests.

5. Discussion and Managerial Implications

5.1. Discussion

This paper provides insight into newsvendor decision making by (i) proposing a method to approximate mismatch costs resulting from level and adjustment behavior and (ii) studying how censored demand, a situation arguably faced in most newsvendor settings, affects level and adjustment behavior and costs. Selected results are summarized in Table 6 (together with results mentioned in §4.3).

By providing a method to approximate portions of expected mismatch costs related to behavioral patterns, we can disambiguate their respective effects on performance. The fact that estimated adjustment cost significantly exceeds the corresponding level cost in three of the four conditions highlights its importance as a factor in understanding the impact of behavioral patterns on the newsvendor model. In light of this, we would like to emphasize that one of the contributions of this paper is the demonstration of how different aspects of the newsvendor model, a rather complex managerial decision setting, result in an intricate combination of behavioral deviations from the normative solution. This suggests that, in addition to understanding the cognitive causes of behavioral patterns in such settings, it is important to understand ways in which level behavior and adjustment behavior coexist.

Table 6 Summary of Selected Results

<table>
<tr><th>Test</th><th>LU</th><th>LC</th><th>HU</th><th>HC</th></tr>
<tr><td>Within pull-to-center region</td><td>***</td><td></td><td>***</td><td></td></tr>
<tr><td>Below pull-to-center region</td><td></td><td>**</td><td></td><td>(*)</td></tr>
<tr><td>Comparing average quantities</td><td colspan="2">>>></td><td colspan="2">>>></td></tr>
<tr><td>Comparing quantity standard deviation</td><td colspan="2">>>></td><td colspan="2">>>></td></tr>
<tr><td>Positive relationship sales standard deviation and quantity standard deviation</td><td>NA</td><td>***</td><td>NA</td><td>***</td></tr>
<tr><td>Positive relationship negative mismatch and update</td><td>***</td><td>***</td><td>***</td><td>***</td></tr>
<tr><td>Positive relationship positive mismatch and update</td><td>***</td><td>***</td><td>***</td><td>***</td></tr>
<tr><td>Comparing level cost</td><td colspan="2">(>)</td><td colspan="2"><<<</td></tr>
<tr><td>Comparing adjustment cost</td><td colspan="2">>>></td><td colspan="2">>>></td></tr>
<tr><td>Comparing behavioral cost</td><td colspan="2">>>></td><td colspan="2"><<<</td></tr>
<tr><td>Learning to avoid level bias</td><td>(*)</td><td>(*)</td><td>***</td><td>***</td></tr>
<tr><td>Learning to avoid adjustment bias</td><td>***</td><td>***</td><td>***</td><td>***</td></tr>
</table>

Notes. "*" significance, "<"/">" significantly smaller/larger, with one, two, and three symbols indicating significance at levels $p = 0.10$, $p = 0.05$, and $p = 0.01$, respectively, and a symbol in parentheses indicating directional support (but not statistically significant).

We also explore the effects of demand censoring on level and adjustment behaviors vis-á-vis the corresponding measures for uncensored demand. Despite subjects having full knowledge of the demand distribution in both LMG and HMG settings, censored demand leads to significantly lower order quantities than uncensored demand does, with the results being robust to settings in which the demand distribution is unknown. This demand-censoring effect led to a significant increase, relative to the uncensored setting, of level costs in the HMG setting equal to 175.0% (34.0% of the normative mismatch cost) and a nonsignificant reduction, relative to the uncensored setting, in level costs in the LMG setting equal to 20.7% (1.8% of the normative mismatch cost). In terms of adjustment behavior, we found that censoring suppresses subjects' order variability relative to uncensored settings, and we conjectured that the reduction in observed variability from censoring is a significant driver of this reduced adjustment behavioral. Our results support the conjecture that the magnitude of adjustments is at least partly driven by the degree of variability that subjects observe. As a result, demand censoring led to a significantly lower adjustment cost (by 35.7%) than that observed with uncensored demand feedback (9.6% of normative mismatch cost) in the HMG setting and a significantly lower adjustment cost (by 49.1%) than that observed with uncensored demand feedback (10.8% of normative mismatch costs) in the LMG setting. The overall effect of demand censoring was a significant increase in behavioral costs, relative to the uncensored setting, in HMG of 52.7% (24.4% of mismatch costs) and a significant decrease of behavioral costs, relative to the uncensored setting, in LMG of 41.1% (12.6% of mismatch costs). These results indicate that more information through full demand feedback can actually lead to worse performance caused by behavioral tendencies.

5.2. Managerial Implications

This research provides managers with insight into how both order adjustment and order level affect behavioral mismatch costs. Furthermore, our results provide guidance as to which of these sources of behavioral cost managers are likely to be most exposed to, given the per-unit profitability of the product they manage—LMG or HMG—and given the setting under which they operate—uncensored or censored demand feedback. As adjustment behavior is addressed by improving the consistency with which an order policy is applied and as order level is addressed by improving the accuracy of that policy with respect to the normative quantity, our results can help managers prioritize improvement efforts. The results suggest that managers in LMG settings tend to have greater exposure to adjustment costs regardless of the demand censoring setting. Managers in HMG uncensored settings would also tend to have greater exposure to adjustment costs, whereas managers in HMG censored settings would tend to be more exposed to level costs. Improvements to the consistency with which an order policy is applied, and its accuracy relative to the normative quantity could be prioritized accordingly.

In general, changing the decision-making process for order quantity preferably comes by building awareness of level and adjustment costs and their sources. Fiske and Taylor (1991) found that salience effects, the posited mechanism driving the difference in observation bias between uncensored and censored settings, can be overcome through training and coaching. Furthermore, Bolton and Katok (2008) found that personal experience led to the greatest improvement in performance in their study. In this vein, the research presented here has also resulted in a pedagogical case, Ludo (A) and (B) (Rudi and Drake 2009, Rudi 2009). The case describes a setting that involves LMG, HMG, and censored demand and introduces the newsvendor concept. After completing a series of periods through a Web-based game, students are led in a discussion of level and adjustment costs, using the decisions they made. At the time of this writing, this case has been adopted for MBA core operations classes, executive MBA programs, and executive education programs, training more than 4,000 managers and executives in the key concepts of the newsvendor model and its central behavioral issues.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2013.1825.

Acknowledgments

Eirill Bø facilitated access to subjects for the experiment, Stig Palmquist programmed the Web-based game, and Agnese Vitali provided assistance for statistical tests. This paper benefited from discussions with Neil Bearden, Philippe Delquie, Serguei Netessine, Anne Grete Preus, Taylor Randall, and Arnstein Aassve. This research was partially funded by the INSEAD Alumni Fund.

References

Benzion U, Cohen Y, Peled R, Shavit T (2008) Decision-making and the newsvendor problem: An experimental study. *J. Oper. Res. Soc.* 59:1281–1287.

Bolton GE, Katok E (2008) Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing Service Oper. Management* 10(3):519–538.

Bolton GE, Ockenfels A, Thonemann UW (2012) Managers and students as newsvendors. *Management Sci.* 58(12):2225–2233.

Bostian AA, Holt CA, Smith AM (2008) Newsvendor "pull-to-center" effect: Adaptive learning in a laboratory experiment. *Manufacturing Service Oper. Management* 10(4):590–608.

Cachon G, Terwiesch C (2009) *Matching Supply with Demand: An Introduction to Operations Management*, 2nd ed. (McGraw-Hill/Irwin, New York).

Ding X, Puterman ML, Bisi A (2002) The censored newsvendor and the optimal acquisition of information. *Oper. Res.* 50(3):517–527.

Feiler DC, Tong JD, Larrick RP (2013) Biased judgment in censored environments. *Management Sci.* 59(3):573–591.

Fiske ST, Taylor SE (1991) *Social Cognition* (McGraw-Hill, Singapore).

Harpaz G, Lee WY, Winkler RL (1982) Learning, experimentation, and the optimal output decisions of a competitive firm. *Management Sci.* 28(6):589–603.

Kremer M, Minner S, Van Wassenhove LN (2010) Do random errors explain newsvendor behavior? *Manufacturing Service Oper. Management* 12(4):673–681.

Lariviere MA, Porteus EL (1999) Stalking information: Bayesian inventory management with unobserved lost sales. *Management Sci.* 45(3):346–363.

Lewis DK (1969) *Convention: A Philosophical Study* (Harvard University Press, Cambridge, MA).

Lurie NH, Swaminathan JM (2009) Is timely information always better? The effect of feedback frequency on decision making. *Organ. Behav. Human Decision Processes* 108(2):315–329.

Massey C, Wu G (2005) Detecting regime shifts: The causes of under- and overreaction. *Management Sci.* 51(6):932–947.

Nahmias S (1994) Demand estimation in lost sales inventory systems. *Naval Res. Logist. Quart.* 41(6):739–757.

Rudi N (2009) Ludo (B). Newsvendor calculations. INSEAD Case, INSEAD, Singapore.

Rudi N, Drake DF (2009) Ludo (A): Ludo Press and the newsvendor problem. INSEAD Case, INSEAD, Singapore.

Schweitzer ME, Cachon GP (2000) Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Sci.* 46(3):404–420.

Su X (2008) Bounded rationality in newsvendor models. *Manufacturing Service Oper. Management* 10(4):566–589.

Taylor SE, Crocker J, Fiske ST, Sprinzen M, Winkler J (1979) The generalizability of salience effects. *J. Personality Soc. Psych.* 37(3):357–368.

Tversky A, Kahneman D (1973) Availability: A heuristic for judging frequency and probability. *Cognitive Psych.* 5(2):207–232.

Tversky A, Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science* 185(4157):1124–1131.