Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence

Maurice E. Schweitzer • Gérard P. Cachon
Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104
Fuqua School of Business, Duke University, Durham, North Carolina 27708
maurice@grace.wharton.upenn.edu • gpc@mail.duke.edu • www.duke.edu/~gpc

In the newsvendor problem a decision maker orders inventory before a one period selling season with stochastic demand. If too much is ordered, stock is left over at the end of the period, whereas if too little is ordered, sales are lost. The expected profit-maximizing order quantity is well known, but little is known about how managers actually make these decisions. We describe two experiments that investigate newsvendor decisions across different profit conditions. Results from these studies demonstrate that choices systematically deviate from those that maximize expected profit. Subjects order too few of high-profit products and too many of low-profit products. These results are not consistent with risk-aversion, risk-seeking preferences, Prospect Theory preferences, waste aversion, stockout aversion, or the consequences of underestimating opportunity costs. Two explanations are consistent with the data. One, subjects behave as if their utility function incorporates a preference to reduce ex-post inventory error, the absolute difference between the chosen quantity and realized demand. Two, subjects suffer from the anchoring and insufficient adjustment bias. Feedback and training did not mitigate inventory order errors. We suggest techniques to improve decision making.
(*Behavioral Operations*; *Newsvendor Inventory Decisions*; *Decision Bias*; *Anchoring*; *Minimizing Ex-Post Inventory Error*)

1. Introduction

In the newsvendor problem a manager sells a product during a short selling season with stochastic demand. The manager has one opportunity to order inventory before the selling season, and no further replenishments are possible. If the order quantity is greater than realized demand, the manager must dispose of the remaining stock at a loss. If the order quantity is lower than realized demand, the manager forgoes some profit. Therefore, in choosing an order quantity the manager must balance the costs of ordering too little against the costs of ordering too much.

The newsvendor problem applies in a broad array of settings. For example, fashion apparel retailers often must submit orders well in advance of a selling season, without opportunity for replenishment during the season. A manufacturer might need to choose its capacity (i.e., its order quantity) before the launch of a new product, knowing that the new product will become obsolete quickly (e.g., computers or cellular phones). Special promotions usually present a similar problem: order too little and the retailer faces irate customers, but order too much and the retailer incurs additional inventory holding costs as it slowly sells the excess inventory. The newsvendor model also applies to individual choice problems, such as health care financing and insurance purchasing (Rosenfield 1986, Eeckhoudt et al. 1991).

There is plenty of anecdotal evidence to suggest that

0025-1909/00/4603/0404$05.00
1526-5501 electronic ISSN

newsvendor decisions can have significant consequences. In the fast-moving computer business it is not uncommon for firms to experience a substantial mismatch between supply and demand: In one year International Business Machines produced $700 million of excess inventory of their ValuePoint line, but in another year they under produced their Aptiva® PC line, and lost potential revenues of more than $100 million (Ziegler 1994, 1995). In 1996, Burger King® restaurants planned to give purchasers of a kids meal a free toy associated with the movie *Toy Story*, but most restaurant owners underestimated demand. Many parents were annoyed since they had to cope with disappointed children (Beatty 1996). Though suggestive of the difficulty managers face in making these types of decisions, these examples provide no indication as to whether managers made good decisions. The optimal order quantity ex-ante is rarely the best order quantity ex-post, and these examples may only represent extreme realizations of demand, rather than biased decision making.

Fisher and Raman (1996) provide some evidence to indicate that managers' decisions do not correspond to the expected profit-maximizing order quantity. They designed an algorithm to apply the newsvendor problem to a fashion apparel manufacturer. This rigorous approach increased profits by about 60% relative to the unassisted decisions made by the firm's managers. They found that managers ordered quantities that were systematically lower than their algorithm's recommendations, but they do not provide an explanation for this bias, nor is it possible to determine if this bias would persist in other settings.

There is some experimental work that has investigated inventory decision making. In one study involving production scheduling the newsvendor problem was a subproblem of the decision task (Carlson and O'Keefe 1969). The authors found that subjects made "almost every kind of mistake" (p. 483). In other studies, Sterman (1989) and Diehl and Sterman (1995) found the anchoring and insufficient adjustment bias in an inventory distribution system experiment with multiple actors, time periods, feedback, and time delay. Croson and Donohue (1998) studied a related problem and found that the design of a supply chain influenced the variability of inventory orders. None of these studies, however, was designed to disentangle biases in the newsvendor context.

This paper seeks to describe and explain managers' newsvendor decisions. These decisions may systematically deviate from profit maximization for several reasons. First, a decision maker may have preferences other than profit maximization. For example, a risk-averse decision maker will systematically order less than the profit-maximizing order quantity (Eeckhoudt et al. 1995). Second, a decision maker may apply a heuristic to choose an inventory level. We consider the anchoring and insufficient adjustment heuristic which assumes a decision maker focuses on a focal value, such as mean demand, and then insufficiently adjusts towards a second value, such as the profit-maximizing order quantity. Third, a decision maker may have a biased forecast of the demand distribution. In this paper, however, we assume that the decision maker knows the distribution of demand. This is a reasonable assumption when the decision maker has access to a substantial amount of historical data for similar products. As Fisher and Raman (1996) discovered, even though a fashion apparel manufacturer changed styles each year, the demand distributions for similar styles closely resembled each other across years. Similar observations have been made at L.L. Bean® (Schleifer 1993).[1]

We conducted two experiments to investigate inventory decisions. We found a consistent too low / too high pattern of orders: Orders for high-profit products were lower than the expected profit-maximizing quantities, while orders for low-profit products were higher than the expected profit-maximizing quantities. (In §2 we define what we mean by high- and low-profit products.) We show that this pattern of choices is not consistent with risk-aversion, risk-seeking preferences, Prospect Theory preferences, loss aversion, waste aversion, stockout aversion, or under-valuing opportunity costs. This pattern, however, is consistent with a preference to reduce ex-post inven-

[1] See Cachon and Schweitzer (1998) for a discussion of newsvendor decision making when the demand distribution is unknown.

tory error and the anchoring and insufficient adjustment heuristic.

2. Descriptive Models of Newsvendor Decision Making

There are several reasons why a decision maker may order an inventory quantity that differs from the expected profit-maximizing quantity. In this section we define the newsvendor problem, present the expected profit-maximizing solution, and describe utility functions and heuristics that could influence the inventory decision process.

2.1. The Newsvendor Problem

In the newsvendor problem a decision maker chooses an order quantity, q, which arrives before the start of a single selling period. Let D be stochastic demand during this period and let μ be its mean. Let F be the distribution function of demand and f the density function. For simplicity, assume F is continuous, differentiable and strictly increasing. Further, assume that the decision maker has an unbiased forecast of the demand distribution and knows F. The decision maker purchases each unit for cost c and sells each unit at price $p > c$. When $q > D$, each unit remaining at the end of the period can be salvaged for $s < c$.

Let $\pi(q, D)$ be realized profit, where

$$\pi(q, D) = (p - s)\min(q, D) - (c - s)q,$$

and expected profit is

$$E[\pi(q, D)] = (1 - F(q))\pi(q, q) + \int_0^q f(x)\pi(q, x)dx.$$

Let $q_n = \arg\max E[\pi(q, D)]$. It is well known that q_n is the unique solution to

$$F(q_n) = \frac{p - c}{p - s}. \tag{1}$$

The ratio $(p - c)/(p - s)$ is called the *critical fractile*, and we use this fractile to classify products. We define a product as a *high-profit* product when

$$\frac{1}{2} \leq \frac{p - c}{p - s} \tag{2}$$

and as a *low-profit* product otherwise. Typical examples of high-profit products include books, bicycles, and fashion apparel. Low-profit products have small margins and low salvage values, and include products such as computers. Throughout this paper we will assume a symmetric demand distribution, so median demand equals mean demand, i.e., $F(\mu) = \frac{1}{2}$. Given that assumption, comparison of (1) and (2) reveals that the expected profit-maximizing order quantity is greater than mean demand for high-profit products, $q_n > \mu$, and less than mean demand for low-profit products, $q_n < \mu$.

2.2. Utility Maximizing Orders

In this section we describe several alternative utility functions. Let w_0 be a decision maker's initial wealth and let $u(w)$ be the decision maker's utility over final wealth. Expected utility is

$$E[u(w_0 + \pi(q, D))] = (1 - F(q))u(w_0 + \pi(q, q)) + \int_0^q f(x)u(w_0 + \pi(q, x))dx. \tag{3}$$

2.2.1. Risk Neutral Preferences. Consider a risk neutral decision maker with utility $u_n(w) = w$. Maximizing $u_n(w)$ is equivalent to maximizing expected profits, $E[\pi(q, D)]$, and q_n is the optimal risk-neutral order quantity.

2.2.2. Risk-Averse and Risk-Seeking Preferences. Consider a risk-averse decision maker with utility $u_a(w)$ who orders q_a, and a risk-seeking decision maker with utility $u_s(w)$ who orders q_s. By definition, $u_a''(w) < 0$ and $u_s''(w) > 0$. Eeckhoudt et al. (1995) demonstrated that a risk-averse decision maker orders less than the normative benchmark, $q_a < q_n$, and a risk-seeking decision maker orders more, $q_s > q_n$.

2.2.3. Prospect Theory Preferences. Alternatively, a decision maker may have reference dependent preferences consistent with Prospect Theory (Kahneman and Tversky 1979). In this case a decision maker will be risk-averse over the domain of gains and risk-seeking over the domain of losses. In the context of the newsvendor problem it is natural to assume the reference point equals current wealth. Consequently, when all possible outcomes are gains, the decision maker will order less than q_n, and when all possible outcomes are losses, the decision maker will order more than q_n. When both gains and losses are possible the utility-maximizing order quantity may either be greater than or less than q_n.[2]

2.2.4. Loss-Averse Preferences. A decision maker may also have preferences for decision attributes other than risk. Consider the following utility function that exhibits loss aversion,

$$u_l(w) = \begin{cases} w & w \geq w_0 \\ \lambda w & w < w_0 \end{cases}, \tag{4}$$

where $\lambda > 1$.[3] Let $\bar{d}(q)$ be the decision maker's break-even sales for a given order quantity,

$$\pi(q, \bar{d}(q)) = 0.$$

Hence,

$$E[u_l(q, D, w_0)] = w_0 + E[u_n(\pi(q, D))] + (\lambda - 1)\int_0^{\bar{d}(q)} f(x)\pi(q, x)dx. \tag{5}$$

Note that the latter term above is negative since $\pi(q, x) < 0$ for $x < \bar{d}(q)$. Let q_l maximize (5).

Theorem 1. *A loss-averse decision maker with utility function* (4) *orders less than the profit-maximizing quantity, $q_l < q_n$. Further, the optimal order quantity is decreasing in the degree of loss aversion, $\partial q_l / \partial \lambda < 0$.*

Proof. Differentiate $E[u_l(w_0 + \pi(q, D))]$ to confirm that it is strictly concave. Note that

$$\frac{dE[u_l(w_0 + \pi(q_n, D))]}{dq} = -(\lambda - 1)(c - s)F(\bar{d}(q_n)) < 0,$$

so $q_l < q_n$ immediately follows. From the implicit function theorem,

$$\frac{\partial q_l}{\partial \lambda} = \frac{-(c - s)F(\bar{d}(q_n))}{-\dfrac{d^2E[u_n(\pi(q_n, D))]}{dq^2}} < 0. \quad \square$$

2.2.5. Waste-Averse Preferences. We assume that a waste-averse decision maker particularly dislikes salvaging excess inventory (Arkes 1996).[4] We model this utility, $u_t(w)$, by assuming the decision maker incurs an additional penalty $t > 0$ for each unit of inventory that must be salvaged at the end of the season:

$$E[u_t(q, D, w_0)] = w_0 + E[u_n(\pi(q, D))] - t\int_0^q f(x)(q - x)dx. \tag{6}$$

Let q_t maximize (6).

Theorem 2. *A waste-averse decision maker orders less than the profit-maximizing quantity, $q_t < q_n$.*

Proof. Differentiate $E[u_t(q, D, w_0)]$ to confirm that it is strictly concave. Note that

$$\frac{dE[u_t(q_n, D, w_0)]}{dq} = -tF(q_n) < 0,$$

so $q_t < q_n$ immediately follows. $\square$

2.2.6. Stockout-Averse Preferences. Alternatively, a decision maker could be averse to losing potential sales. We refer to these preferences as stockout-aversion. For example, stocking out could lead to irate customers or a loss in market share. Let $u_m(q, D, w_0)$ be the decision maker's stockout-averse utility,

$$E[u_m(q, D, w_0)] = w_0 + E[u_n(\pi(q, D))] - \alpha\int_q^\infty f(x)(x - q)dx, \tag{7}$$

where α is a positive constant. Let q_m maximize (7).

[2] We do not demonstrate this result formally since it is straightforward to construct an example in which this holds.

[3] The loss-aversion preference function is a special case of a Prospect Theory preference function. If a decision maker exhibits loss aversion with diminishing sensitivity, the utility function will be strictly convex for losses and the subsequent results do not apply.

[4] Arkes (1996) documents preferences to "avoid the appearance of wastefulness" (p. 222) even when such behavior compromises other goals. For example, a person who is not hungry may finish leftovers to avoid appearing wasteful. In this paper we define such preferences as waste averse.

Theorem 3. *A stockout-averse decision maker orders more than the profit-maximizing quantity,* $q_m > q_n$.

Proof. Differentiate $E[u_m(q, D, w_0)]$ to confirm that it is strictly concave. Note that

$$\frac{dE[u_m(q_n, D, w_0)]}{dq} = \alpha(1 - F(q_n)) > 0,$$

so $q_m > q_n$ immediately follows. □

2.2.7. Underestimated Opportunity Costs. Instead of stockout aversion a decision maker may undervalue forgone sales. For example, when $D > q$, opportunity costs equal

$$\pi(D, D) - \pi(q, D) = (p - c)(D - q),$$

which is the difference between potential profits and actual profits. If the decision maker discounts the marginal value of increased sales, $(p - c)$, opportunity costs will be undervalued.

Using integration by parts, the risk-neutral utility function can be written as

$$u_n(q, D, w_0) = w_0 + (p - c)q - ((p - c) - (c - s)) \int_0^q F(x)dx.$$

A natural way to model underestimating opportunity costs is to assume a decision maker underestimates the two $(p - c)$ terms. Let $u_0(q, D, w_0)$ be the utility function of a decision maker that underestimates opportunity costs,

$$u_o(q, D, w_0) = w_0 + \beta(p - c)q - (\beta(p - c) - (c - s)) \int_0^q F(x)dx,$$

where $0 \leq \beta < 1$. Let q_o be the order quantity that maximizes $u_o(q, D, w_0)$.

Theorem 4. *When a decision maker undervalues opportunity costs, the decision maker orders less than the profit-maximizing quantity,* $q_o < q_n$.

Proof. Differentiate $u_o(q, D, w_0)$ to confirm that it is strictly concave. Note that

$$\frac{du_o(q_n, D, w_0)}{dq} = -(1 - \beta)\,\frac{(p - c)(c - s)}{p - s} < 0,$$

so it follows immediately that $q_o < q_n$. □

2.2.8. Preference for Minimizing Ex-Post Inventory Error. The final preference function we consider is for a decision maker who cares about reducing ex-post inventory error, the deviation between the order quantity and realized demand. This preference could develop from the decision maker's anticipated regret or disappointment from not choosing the *ex-post* optimal order quantity (realized demand), even though that order quantity is rarely the *ex-ante* optimal order quantity (Bell 1982, 1985).

Let $u_e(q, D, w_0)$ be the ex-post inventory error utility function,

$$u_e(q, D, w_0) = w_0 + u_n(\pi(q, D)) - \delta(|q - D|),$$

where $\delta' > 0$ and $\delta(0) = 0$. The function $\delta(\cdot)$ is the penalty for choosing an order quantity that deviates from realized demand. So

$$E[u_e(q, D, w_0)] = w_0 + E[u_n(\pi(q, D))] - \int_0^\infty f(x)\delta(|q - D|)dx. \tag{8}$$

Let q_e be an order quantity that maximizes (8).

Theorem 5. *If* $u_e(\cdot)$ *is the decision maker's utility function and F is symmetric about its mean* (*i.e.,* $f(\mu + y) = f(\mu - y)$, $y \geq 0$) *then for high-profit products* $\mu \leq q_e \leq q_n$ *and for low-profit products* $\mu > q_e > q_n$.

Proof. Since F is symmetric about its mean and sales are never negative, 2μ is an upper bound on demand, i.e., $F(2\mu) = 1$. Differentiate,

$$\frac{dE[u_e(q, D, w_0)]}{dq} = \frac{dE[u_n(\pi(q, D))]}{dq} - \int_0^{2\mu} f(x)(q - x)\delta'(|q - x|)dx. \tag{9}$$

Since F is symmetric about its mean and bounded,

$$-\int_{0}^{2\mu} f(x)(\mu - x)\delta'(|\mu - x|)dx = 0.$$

From the above, the second term in (9) is negative when $q > \mu$, and positive when $q < \mu$. Since $u_e(q, D, w_0)$ is continuous in q, any q_e must satisfy the first order condition,

$$\frac{dE[u_e(q_e, D, w_0)]}{dq} = 0.$$

(Note that $u_e(q, D, w_0)$ is not necessarily strictly concave in q. Hence, there can exist multiple profit-maximizing order quantities.) Consider a high-profit product, so $q_n \geq \mu$. For any $q > q_n$,

$$\frac{dE[u_e(q, D, w_0)]}{dq} < 0,$$

hence $q_e \leq q_n$. For any $q < \mu$,

$$\frac{dE[u_e(q, D, w_0)]}{dq} > 0,$$

hence $q_e \geq \mu$. The analogous argument demonstrates $\mu > q_e > q_n$ for a low profit product. □

From Theorem 5, for a broad class of distribution functions, a decision maker who cares about ex-post inventory error will make inventory choices that exhibit a *too low/too high* pattern: orders for high profit products are lower than q_n, while orders for low-profit products are higher than q_n. The pattern may also be observed for a decision maker with Prospect Theory preferences, but only if the domain of outcomes includes both gains and losses. This pattern is not consistent with the other utility functions considered.

2.3. Anchoring and Insufficient Adjustment

Instead of maximizing utility to choose an inventory quantity, a decision maker may use a decision heuristic. For example, a decision maker may use anchoring and insufficient adjustment. Anchoring and insufficient adjustment implies that a decision maker selects an available anchor, and then modifies or adjusts away from this quantity (Kahneman et al. 1982). In this context, there are several available anchors including expected demand, a prior order quantity, or a past realization of demand. Sterman (1989) argued that subjects in his study chose order quantities based on current stock levels and some insufficient adjustment toward the desired stock levels.

In our setting, we consider two alternative anchoring and insufficient adjustment heuristics. The first heuristic, which we call the *mean anchor* heuristic, assumes a decision maker anchors on mean demand and adjusts towards the optimal order quantity, q_n. This heuristic predicts the same too low/too high pattern of choices as the preference to reduce ex-post inventory error. The second heuristic, which we call the *chasing demand* heuristic, assumes a decision maker anchors on a prior order quantity and adjusts towards prior demand. This heuristic does not predict initial choices, but does predict adjustment patterns across a series of choices. Further, it makes no formal claim regarding the relationship between mean demand and choices.

2.4. Summary

We consider preference and process explanations for managerial decisions that deviate from the expected profit-maximizing order. Table 1 identifies sources of deviations and describes their consequences. Note that these sources of deviations need not be mutually exclusive. Managers may be influenced by multiple sources, and the relative strength of different effects is likely to be moderated by task and contextual factors.

3. Newsvendor Experiments

We explore newsvendor decision making in two experiments across different profit-margin conditions. These experiments are designed to disentangle alternative explanations for observed deviations from the expected profit-maximizing order quantity.

3.1. Experiment 1: Uniform Distribution Experiment

Our first experiment investigates inventory orders in a repeated environment. The demand distribution is known and subjects made decisions for both high- and low-profit margin products.

3.1.1. Methods. We recruited 34 subjects from a Duke University M.B.A. operations management course. Each subject received a self-starting computer

Table 1 Source of Deviations from the Expected Profit-Maximizing Order Quantity

	Order quantity relative to the expected profit-maximizing order quantity*	
	High Profit	Low Profit
Preferences:		
Risk averse	−	−
Risk seeking	+	+
Prospect Theory	+/−	+/−
Underestimate opportunity costs	−	−
Waste averse	−	−
Stockout averse	+	+
Minimize ex-post inventory error	−	+
Anchoring and insufficient adjustment heuristic:		
Anchor on mean demand, Adjust toward optimal order quantity	−	+
Anchor on prior order quantity, Adjust toward prior demand	+/−	+/−

* "−" = lower quantity; "+" = higher quantity; "+/−" = either higher or lower quantity.

program that described an inventory problem for selling "wodgets." Subjects were not informed of the total number of rounds or the price and cost of wodgets in future rounds. The computer program prompted subjects to make 30 inventory decisions. Profits were calculated in francs, and subjects were informed that francs would be converted to cash at an exchange rate of 300 francs to the dollar and that one subject would be selected at random and paid in dollars.

Subjects were provided with cost and demand data for each decision. Wodgets were sold for 12 francs and salvaged for 0 francs. In the low-profit condition wodgets were purchased for 9 francs and in the high-profit condition wodgets were purchased for 3 francs. Therefore, the critical fractiles were 75% = (12-3)/(12-0) in the high-profit condition, and 25% = (12-9)/(12-0) in the low-profit condition. Demand was discrete and uniformly distributed between 1 and 300, and the corresponding expected profit-maximizing order quantities were q_n = 225 in the high-profit condition and q_n = 75 in the low-profit condition. These order quantities correspond to the 75% fractile (0.75*300) and the 25% fractile (0.25*300).

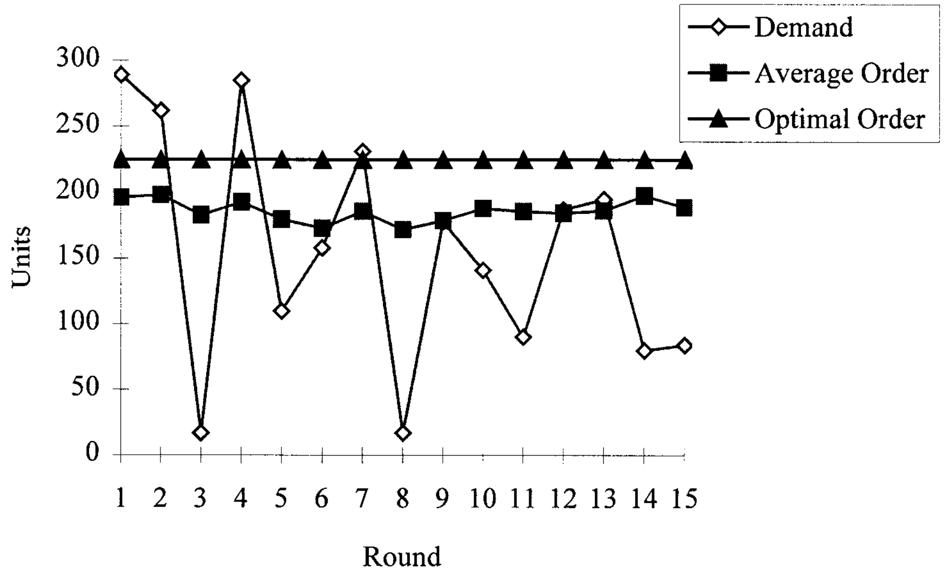
Before making a decision, subjects could solicit the following information for any order quantity: The profit distribution (both as a table and as a graph) the probability sales will exceed the order, the probability sales will be lower than the order, the break-even sales level, and the probability sales will be no lower than the break-even sales level. They could also view their results from prior rounds in a table that displayed their past order quantities, realizations of demand, profit from each round, and cumulative profit. This information was presented in both tables and graphs. After making an inventory decision, subjects learned actual demand for the round, and their realized profit or loss. After each round, subjects viewed an updated table of results that included their order quantity, actual demand, profit, and cumulative profit. The vector of actual demand was randomly determined prior to the experiment and was the same for every subject. Demand for each round for each profit condition is depicted in Figures 1 and 2.

Each subject made 15 inventory decisions under both a high-profit and a low-profit condition. The order of the high- and low-profit conditions was randomly determined by the computer program; 20 subjects faced the high-profit condition first, and 14 subjects faced the low-profit condition first.

3.1.2. Results. A total of 34 subjects completed the experiment. A computer error rendered the data from one subject unusable; as a result, we exclude this subject from analysis, and report results from 33 subjects.

We first examine subjects' initial inventory decisions. Analysis of these decisions enable us to make a between-subjects comparison of treatment effects that is unconfounded by the potential effects of experience, feedback, and treatment condition order. Inventory orders were above mean demand (150) in the high-profit condition and below mean demand in the low-profit condition. On average, the first inventory order of subjects who faced the high-profit condition first was significantly higher than the first inventory order of subjects who faced the low-profit condition

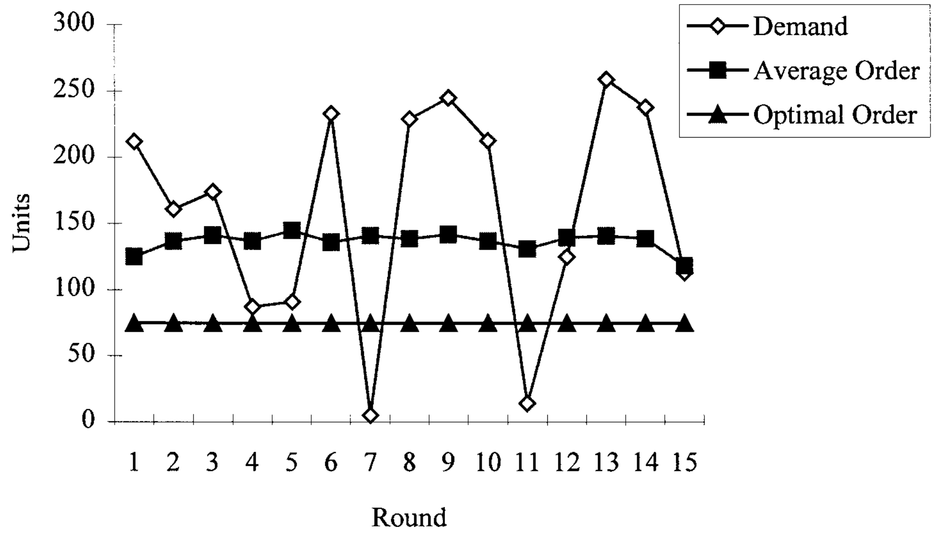
Figure 1 Demand and Order Quantities for the High-Profit Condition in Experiment 1



first: 178.25 units and 139.62 units, respectively (a between subjects t-test yielded $p < 0.01$). Consistent with the too low/too high pattern, subjects' average first inventory orders were less than the expected profit-maximizing quantity of 225 in the high-profit condition, and greater than the expected profit-maximizing quantity of 75 in the low-profit condition; these differences were statistically significant ($t(19) = 4.96$, $p < 0.0001$ and $t(13) = 10.71$, $p < 0.0001$, respectively).

The average order quantities across all inventory decisions also exhibited the too low/too high pattern. We calculated each subject's average order quantity in the high- and low-profit conditions. The average high-

Figure 2 Demand and Order Quantities for the Low-Profit Condition in Experiment 1



profit order, 176.68, was significantly lower than the expected profit-maximizing order quantity of 225 ($t(32) = 6.58$, $p < 0.001$), and the average low-profit order, 134.06, was significantly higher than the expected profit-maximizing order quantity of 75 ($t(32) = 12.15$, $p < 0.001$). In addition, we found that inventory orders were higher when the low-profit condition was presented first, but this difference was not significant ($F(1, 31) = 0.98$, $p = n.s.$).

To analyze the 15 inventory orders under each of the two profit conditions for each subject we conducted a double repeated measures generalized linear model. As expected, we found a significant profit effect. Subjects ordered higher amounts under the high-profit condition than they did under the low-profit condition across both order conditions (Wilks' Lambda = 0.5984; $F(1, 31) = 20.8048$, $p < 0.0001$). More importantly, subjects did not significantly change their inventory orders over time (Wilks' Lambda = 0.4735; $F(14, 18) = 1.4296$, $p = n.s.$). That is, there was no main effect for round. This result is also seen in Figures 1 and 2, which depict average order quantities and actual demand for each round. Finally, we found no significant interaction between presentation order and the profit condition.

We next considered the mean anchor heuristic. Recall that this heuristic assumes decision makers anchor on expected demand, 150, and adjust towards the normative order quantity, 225 in the high-profit case and 75 in the low-profit case. We define q_n adjustment scores to be $(q - \mu)/(q_n - \mu)$ in the high-profit condition and $(\mu - q)/(\mu - q_n)$ in the low-profit condition, where q is a subject's order quantity. We calculated the average q_n adjustment scores for each subject in the high- and low-profit cases. Across subjects average q_n adjustment scores were higher in the high-profit case than in the low-profit case, 0.36 versus 0.21, but the difference between these scores was not significant, ($t(64) = 1.22$, $p = n.s.$). That is, subjects' inventory orders were slightly closer to the normative order quantity in the high-profit case (about a third of the way between mean demand and the normative order quantity) than they were in the low-profit case (less than a quarter of the way between mean demand and the normative order quantity).

We also examined adjustments for the chasing demand heuristic. We analyzed 28 rounds of inventory decisions for each subject, excluding the first and sixteenth rounds when subjects encountered a new profit margin condition. Overall, most decisions (64.3% or 594 decisions) were characterized by repeat choice. When subjects did change their order quantity across rounds, they were more than twice as likely to adjust their order quantity in the direction of prior demand (24.7% or 228 decisions) than away from it (11.0% or 102 decisions). For each subject we calculated the number of times they adjusted toward and away from prior demand. On average, the number of adjustments toward prior demand exceeded the number of adjustments away from prior demand, 6.94 versus 3.09, $t(64) = 3.42$, $p < 0.001$.

Subjects were also more likely to change their order quantity in early rounds than in late rounds.[5] As described by Table 2, in Rounds 2 through 8 subjects were more likely to change their order quantity than in Rounds 24 through 30. When the high-profit margin condition was presented first, subjects changed their order quantity an average of 3.80 times in Rounds 2 through 8 and 1.85 times in Rounds 24 through 30, $t(38) = 2.99$, $p < 0.01$. When the low-profit margin condition was presented first, subjects changed their order quantity an average of 2.69 times in Rounds 2 through 8 and 1.46 times in Rounds 24 through 30, $t(24) = 1.37$, $p = n.s.$

We define the d_{t-1} adjustment scores to be $(q_t - q_{t-1})/(d_{t-1} - q_{t-1})$, where q_t and d_t are the order quantity and realized demand, respectively, in round t. We found that d_{t-1} adjustment scores were greater towards prior demand than away from prior demand when the high profit margin condition was presented first. For each subject we calculated average adjustment scores toward and away from prior demand. When the high-profit condition was presented first, average adjustment scores toward and away from prior demand were 0.31 and -0.11, respectively, $t(38)$

[5] We found differences in adjustment patterns across presentation-order conditions, $\chi^2(2) = 13.72$, $p < 0.001$, and report results for both presentation-order conditions separately.

Table 2 Percentage of Order Quantity Adjustments and Median d_{t-1} Adjustment Scores across Rounds and Presentation Conditions for Experiment 1*

	Low Profit Condition First			High Profit Condition First		
Rounds	No Change	Toward	Away	No Change	Toward	Away
2–8	61.54%	34.07% (0.30)	4.40% (−4.20)	45.71%	43.57% (0.60)	10.71% (−0.38)
9–15	62.64%	13.19% (0.37)	24.18% (−0.71)	65.00%	27.14% (0.36)	7.86% (−0.85)
17–23	67.03%	14.29% (0.72)	18.68% (−0.86)	64.29%	27.86% (0.33)	7.86% (−0.32)
24–30	79.12%	13.19% (0.91)	7.69% (−0.83)	73.57%	15.71% (0.49)	10.71% (−1.79)
All	67.58%	18.68% (0.42)	13.74% (−0.86)	62.14%	28.57% (0.41)	9.29% (−0.46)

* d_{t-1} adjustment scores are $(q_t - q_{t-1})/(d_{t-1} - q_{t-1})$, where q_t and d_t are the order quantity and demand, respectively, in round t. Median adjustment scores are reported in parentheses.

$= 2.75$, $p < 0.01$. When the low-profit condition was presented first, average adjustment scores toward and away from prior demand were 0.24 and −0.27, respectively, $t(24) = 0.18$, $p = n.s.$

We also examined the relationship between order quantity adjustments and the absolute difference between prior order quantity and prior demand. We divided the absolute differences between prior order quantities and prior demand into quartiles, and report the percentage of adjustments toward prior demand following these differences in Table 3. We also calculated the number of adjustments each subject made toward prior demand following the largest and smallest quartile differences between their prior order quantity and prior demand. When the high-profit margin condition was presented first, subjects adjusted their order quantity in the direction of prior demand an average of 2.55 times when the difference was large and 1.70 times when the difference was small, $t(38) = 1.40$, $p = n.s.$ When the low-profit margin condition was presented first, subjects adjusted their order quantity in the direction of prior demand an average of 1.08 times when the difference was large and 1.62 times when the difference was small, $t(24) = 0.87$, $p = n.s.$

At the individual level, almost every subject ordered more than the expected profit-maximizing 75 units in the first round of the low-profit condition (29 of 33 subjects did). Three subjects ordered exactly 75 units in the first round, but two of these subjects subsequently increased their order quantities. Similarly, most subjects ordered much less than the optimal 225 units in the first round of the high-profit

Table 3 Percentage of Order Quantity Adjustments Following Large and Small Differences Between Prior Order Quantity and Prior Demand for Experiment 1

	Low Profit Condition First			High Profit Condition First		
Absolute Difference Between Prior Order and Prior Demand	No Change	Toward	Away	No Change	Toward	Away
Largest Quartile of Differences	71.8%	16.5%	11.8%	58.8%	37.5%	3.7%
Smallest Quartile of Differences	63.7%	23.1%	13.2%	62.0%	24.8%	13.1%
All Rounds	67.58%	18.68%	13.74%	62.14%	28.57%	9.29%

condition (26 of 33 did). Of the four subjects who ordered exactly 225 units in the first round, two decreased their order quantities in subsequent rounds.

3.1.3. Discussion. Experiment 1 establishes the too low/too high pattern in inventory orders as a robust phenomenon that persisted across multiple periods with feedback. These results cannot be explained exclusively by risk aversion, risk-seeking preferences, loss aversion, waste aversion, stockout aversion, or underestimation of opportunity costs. A combination of these preferences, such as stockout aversion and risk aversion, could lead to the too-low/too-high pattern, provided the "too high" preference is relatively stronger in the low-profit condition and the "too low" preference is relatively stronger in the high profit condition. Prospect Theory preferences, ex-post inventory preferences, or an anchoring and insufficient adjustment process, however, offer a more parsimonious explanation of the data; each of these explanations can account for a too low/too high pattern on its own.

Two additional patterns in the data merit discussion. First, subjects' order quantities were closer to q_n in the high-profit condition than in the low-profit condition. On average, subjects ordered 48 units fewer than q_n in the high-profit condition and 59 units more than q_n in the low-profit condition. This difference is more dramatic when considered in terms of expected profit. Instead of the expected earnings of 675 and 75 francs for ordering q_n in the high- and low-profit conditions, a subject who ordered the average order quantities in this experiment could expect to earn 644 and 29 francs, a 5% and 61% decline. Neither a preference for minimizing ex-post inventory error nor an anchoring and insufficient adjustment heuristic predict this asymmetry. Prospect Theory could explain the asymmetry if subjects were relatively more risk-seeking in losses than risk averse in gains. Alternatively, the asymmetry could be explained with a combination of preferences. For example, suppose a decision maker wishes to minimize ex-post inventory error and avoid stockouts. Recall that the latter preference causes the decision maker to increase their order quantities. Since stockouts are more likely in the low-profit condition than in the high-profit condition, the decision maker is more likely to increase their orders in the low-profit condition than in the high-profit condition, all else being equal. That effect would lower q_n adjustment scores in the low-profit condition relative to the high-profit condition. There is also a third plausible explanation. Since the absolute stakes were higher in the high-profit condition, subjects may have exerted more effort in that condition, leading to choices that were closer to optimal.

The second noteworthy pattern is that the average choice was remarkably stable across rounds, as is clearly observed in Figures 1 and 2. This result suggests that feedback and experience did not help subjects order quantities closer to the optimal order quantity. Further, while there is some evidence of a chasing demand heuristic, the impact of that heuristic is small. Indeed, in most rounds subjects did not adjust their orders, and the adjustment percentages declined with the number of rounds.

3.2. Experiment 2: High-Demand Distribution Experiment

This experiment was designed to disentangle the preference function explanations of the too low/too high pattern. Specifically, we consider Prospect Theory preferences and a preference to reduce ex-post inventory error. Prospect Theory preferences predict the too low/too high pattern only when both positive and negative profits are possible, while ex-post inventory preferences predict the too low/too high pattern across both gain and loss domains. This experiment was similar to Experiment 1, except the demand range was increased so that losses were not possible: a subject would earn a positive profit even if maximum demand was ordered and minimum demand was realized. Since this setting lies entirely in the domain of gains, Prospect Theory predicts subjects will be risk averse and therefore always order less than q_n. Hence, in this setting the too low/too high pattern is not consistent with Prospect Theory.

As before, subjects in this experiment received feedback after each round of decisions. In addition, all of the subjects in this experiment had formal training in the newsvendor problem. Consequently, this experiment investigates whether the too low/too high pattern persists over a domain that includes only gains and is robust to both prior training and feedback.

3.2.1. Methods. The methods employed in this experiment were the same as those in Experiment 1, with two exceptions. First, demand in this version includes two ranges of the uniform distribution, a low range of [1,300] and a high range of [901,1200]. In the high-range condition subjects cannot incur a loss. (Even if subjects order the maximum quantity of demand, 1200, and realize minimum demand, 901, they would earn a positive profit even in the low-profit condition.) Second, subjects in this experiment were randomly assigned by the computer to either the low- or high-demand-range condition. Subjects made all of their inventory decisions over this range. As in Experiment 1, subjects were also randomly assigned to receive either the low- or high-profit condition first, and made two sets of 15 inventory decisions under both low and high profit conditions. This design created four treatment conditions across the two demand-range conditions and the two presentation orders of the profit conditions.

As in the previous two experiments, critical fractiles were 75% and 25% in the high- and low-profit conditions, respectively. In the low-demand range, expected profit-maximizing order quantities were $q_n = 225$ in the high-profit condition and $q_n = 75$ in the low-profit condition. In the high-demand range, expected profit-maximizing order quantities were $q_n = 1125$ in the high-profit condition and $q_n = 975$ in the low-profit condition.

We recruited 44 second year M.B.A. students from Duke University to participate in this experiment. There was no overlap between subjects in this experiment and the previous experiment. Unlike subjects in the prior experiment, subjects in this experiment had completed a course in operations management the previous year that covered the newsvendor problem.[6]

3.2.2. Results. Since the computer randomly assigned subjects to each of these treatments, we did not obtain an equal number of subjects in each cell: eight subjects received the high-demand range, low-profit condition first; sixteen received the high-demand range, high-profit condition first; nine received the low-demand range, low-profit condition first; and eleven received the low-demand range, high-profit condition first.

Across both the low- and the high-demand range treatment conditions the too low/too high pattern characterized inventory orders. Subjects' average order quantities under the high-profit condition for the low-demand range and high-demand range treatment conditions were 186.88 and 1092.55, respectively. These were both significantly below the expected profit-maximizing order quantities of 225 ($t(19) = 4.22$, $p < 0.001$) and 1125 ($t(23) = 3.22$, $p < 0.001$). The average order quantities under the low-profit condition for the low-demand range and high-demand range treatment conditions were 142.17 and 1021.81. These were both significantly above the expected profit-maximizing order quantities of 75 ($t(19) = 7.10$, $p < 0.001$) and 975 ($t(23) = 4.78$, $p < 0.001$). All of these differences are significant in two-sided t-tests at Bonferroni corrected p-value levels of 0.00625 (at a group alpha level equal to 0.05 for 4 tests). The amounts ordered in the low-demand range treatment condition in this experiment were similar to those in Experiment 1; in the high-profit condition average orders were 186.88 versus 176.68 ($t(51) = 0.87$, $p = n.s.$), and in the low-profit condition average orders were 142.17 versus 134.06 ($t(51) = 0.84$, $p = n.s.$).

As before, we compared subjects' first inventory orders. The sample size of 44 inventory orders across the four treatment cells, however, is relatively small. The average first inventory orders under the high-profit condition for the low-demand range and high-demand range treatment conditions were 163.64 and 1059.69, and the average first inventory orders under

[6] The subjects were second year M.B.A. students who had taken the operations management core class in their first year. That course spent one two-hour session on the newsvendor problem. The optimal order quantity was explained and the students participated in a discussion of the L.L. Bean® case (Schleifer 1992) which describes an application of the newsvendor problem. Students were aware that the newsvendor problem would be covered in their final exam.

the low-profit condition for the low-demand range and high-demand range treatment conditions were 159.67 and 1034.38. The high-profit inventory levels were too low relative to the expected profit-maximizing quantities of 225 and 1125 ($t(10) = 3.94$, $p < 0.01$ and $t(15) = 65.03$, $p < 0.0001$, respectively), and the low-profit inventory levels were too high relative to the expected profit-maximizing quantities of 75 and 975 ($t(8) = 143.3$, $p < 0.0001$ and $t(7) = 4.20$, $p < 0.01$, respectively).

We also report results from a doubly repeated measures analysis of variance conducted on the complete set of data. Subjects ordered significantly more under the high-demand range condition than the low-demand range condition; a between-subjects test yielded an $F(1, 41) = 11344.89$, $p < 0.0001$. Subjects ordered more under the high-profit condition than the low-profit condition; a within-subjects test produced a Wilks' Lambda = 0.5701; $F(1, 41) = 30.9131$, $p < 0.0001$. We did not, however, find a significant round effect (Wilks' Lambda = 0.9046; $F(14, 28) = 0.2108$, $p = n.s.$) or any significant interactions between demand range, profit condition, presentation order, and round.

As in Experiment 1, we calculated mean and median q_n adjustment scores and list results for the high-profit margin condition first. In the low-demand range mean adjustment scores were 0.49 and 0.10 and median adjustment scores were 0.67 and 0.0; in the high-demand range mean adjustment scores were 0.57 and 0.38 and median adjustment scores were 0.67 and 0.25. Adjustment scores in the high-profit cases were similar to each other, but high-profit scores were significantly greater than low-profit scores, and low-profit scores in the high range were significantly higher than low-profit scores in the low range. (These differences were significant in multiple t-tests with a Bonferroni corrected p-value of 0.0042, for a group alpha equal to 0.05. We also conducted a multiple median test and found $\chi^2(3) = 47.54$, $p < 0.0001$.)

We next examined d_{t-1} adjustment scores to investigate the operation of a chasing demand heuristic. As before, we examined order quantity changes in Rounds 2 through 15 and 17 through 30, excluding the first and sixteenth rounds when subjects encountered an initial or new profit-margin condition. The order adjustment patterns in this experiment were very similar to those in Experiment 1, and we report results across rounds in Table 4. Across both range and presentation order conditions, subjects were most likely to repeat their order quantity decision, and next most likely to adjust their order quantity in the direction of prior demand. For each subject we compared the number of times they changed their order quantity in Rounds 2 through 8 with the number of times they changed their order quantity in rounds 24 through 30. On average, subjects changed their order quantity more often in early rounds than in late rounds. For the low-range, low-profit first condition, low-range, high-profit first condition, high-range, low-profit first condition, and high-range, high-profit first condition the average number of adjustments were 3.89 versus 1.44 ($t(16) = 2.70$, $p < 0.05$), 2.18 versus 2.27 ($t(20) = 0.10$, $p = n.s.$), 4.50 versus 5.25 ($t(14) = 0.57$, $p = n.s.$), and 3.94 versus 1.94 ($t(30) = 2.42$, $p < 0.05$), respectively.

Subjects were also more likely to change their order in the direction of prior demand when the absolute difference between their prior order quantity and prior demand was large. As in Experiment 1, we divided the absolute differences between prior order quantity and prior demand into quartiles, and compared adjustment rates following the largest and smallest differences. We report these adjustment rates in Table 5. For each subject we computed the number of times they adjusted their order quantity in the direction of prior demand following a large and a small difference between their prior order and prior demand. On average, subjects were more likely to adjust their order quantity in the direction of prior demand following large differences. For the low-range, low-profit first condition, low-range, high-profit first condition, high-range, low-profit first condition, and high-range, high-profit first condition the average number of adjustments were 2.00 versus 0.89 ($t(16) = 1.70$, $p = n.s.$), 2.00 versus 1.00 ($t(20) = 1.48$, $p = n.s.$), 2.75 versus 0.63 ($t(14) = 1.44$, $p = n.s.$), and 2.94 versus 1.69 ($t(30) = 2.19$, $p < 0.05$).

3.2.3. Discussion. We again found the too low/too high pattern in inventory orders. Prospect Theory

Table 4 Percentage of Order Quantity Adjustments and Median d_{t-1} Adjustment Scores Across Rounds and Presentation Conditions for Experiment 2*

		Low Profit First			High Profit First		
Demand Condition	Rounds	No Change	Toward	Away	No Change	Toward	Away
Low Range First	2–8	44.43%	44.43%	11.14%	68.86%	26.00%	5.14%
			(0.53)	(−0.89)		(0.60)	(−0.32)
	9–15	68.19%	22.25%	9.56%	74.00%	16.86%	9.14%
			(0.26)	(−1.08)		(0.47)	(−0.68)
	17–23	82.45%	9.56%	7.99%	67.48%	24.68%	7.85%
			(1.34)	(−0.71)		(0.66)	(−0.40)
	24–30	79.32%	11.13%	9.56%	67.57%	18.14%	14.29%
			(0.20)	(−1.17)		(0.60)	(−1.31)
	All	68.64%	21.82%	9.54%	69.52%	21.39%	9.09%
			(0.42)	(−0.91)		(0.60)	(−0.60)
High Range First	2–8	64.19%	30.39%	5.42%	43.71%	45.57%	10.71%
			(0.78)	(−5.00)		(0.56)	(−1.39)
	9–15	71.43%	25.00%	3.57%	59.00%	33.00%	8.00%
			(0.24)	(−0.32)		(0.59)	(−0.68)
	17–23	69.61%	30.39%	0.00%	62.48%	31.24%	6.28%
			(0.21)	(−)		(0.70)	(−0.69)
	24–30	75.00%	25.00%	0.00%	72.29%	18.71%	9.00%
			(0.08)	(−)		(0.50)	(−1.41)
	All	70.08%	27.67%	2.25%	59.37%	32.13%	8.50%
			(0.26)	(−1.00)		(0.58)	(−0.89)

* d_{t-1} adjustment scores are $(q_t - q_{t-1})/(d_{t-1} - q_{t-1})$, where q_t and d_t are the order quantity and demand, respectively, in round t. Median adjustment scores are reported in parentheses.

cannot explain this pattern in the high-demand range condition, since it predicts subjects will order less than the profit-maximizing quantity in the domain of gains. As before, this pattern was robust across multiple periods with feedback, even among subjects who had received prior training in the newsvendor problem. There are several possible explanations. First, the typical M.B.A. training in the newsvendor problem is

Table 5 Percentage of Order Quantity Adjustments Following Large and Small Differences Between Prior Order Quantity and Prior Demand for Experiment 2

		Low Profit First			High Profit First		
Demand Condition		No Change	Toward	Away	No Change	Toward	Away
Low Range First	Largest Quartile of Differences	67.7%	27.7%	4.6%	65.9%	26.8%	7.3%
	Smallest Quartile of Differences	70.2%	11.9%	17.9%	71.6%	13.6%	14.8%
	All Rounds	68.64%	21.82%	9.54%	69.52%	21.39%	9.09%
High Range First	Largest Quartile of Differences	61.0%	37.3%	1.7%	57.1%	39.5%	3.4%
	Smallest Quartile of Differences	71.2%	16.9%	11.9%	59.6%	23.7%	16.7%
	All Rounds	70.08%	27.67%	2.25%	59.37%	32.13%	8.50%

not sufficient to ensure that subjects will choose the optimal order quantity. However, since there was a significant amount of elapsed time between when the subjects received the training (in their first year of classes) and when they participated in the experiment (in their second year), it is possible that the training was effective for some period of time (e.g., up to the final exam) and then the subjects began to forget their training. Alternatively, these decision makers may have been able to evaluate the optimal order, but nevertheless chose to order a different quantity due to other considerations, such as a desire to reduce ex-post inventory error.

As in Experiment 1, we again find weak support for the chasing demand heuristic. Subjects were more likely to adjust their order quantities toward prior demand than away from prior demand. Subjects were also more likely to adjust their order quantities in early rounds, and were more likely to adjust their order quantity toward prior demand when the difference between their prior order quantity and prior demand was large. Most of the time, however, subjects did not adjust their decisions, and across rounds the average order quantity was relatively stable.

4. General Discussion

We found that choices systematically deviated from those that would maximize expected profits. Subjects consistently ordered amounts lower than the expected profit-maximizing quantity for high-profit products and higher than the expected profit-maximizing quantity for low-profit products. This too low/too high pattern of choices cannot be explained by risk aversion, risk-seeking preferences, loss avoidance, waste aversion, or underestimating opportunity costs. Preferences consistent with Prospect Theory (risk aversion over gains and risk seeking over losses) can explain some, but not all, of the data in our experiments. Instead, inventory orders in these studies were consistent with a preference to reduce ex-post inventory error, the absolute deviation between the amount ordered and realized demand. This preference causes subjects to choose order quantities, which are too close to mean demand thereby leading them to order too little of high-profit products (for which the optimal order is greater than mean demand) and too much of low-profit products (for which the optimal order is less than mean demand). These data are also consistent with an anchoring and insufficient adjustment process in which subjects anchor on mean demand and insufficiently adjust toward optimal order quantities.

We find weak support for a chasing demand heuristic. When subjects adjusted their order quantity from the prior round, they were more likely to adjust their order quantity toward prior demand than away from prior demand. Overall, however, the chasing demand heuristic cannot account for the patterns observed in these experiments. In particular it does not explain initial decisions, nor does it account for the relative stability of choices across rounds.

This too low/too high pattern of choices was not symmetric across low and high profit conditions. Orders were closer to expected demand for low profit products than for high profit products. This result raises several questions, and further work remains to understand this result. Possible explanations include the presence of stockout aversion in addition to the ex-post inventory error preference. Alternatively, the anchoring and insufficient adjustment process may be context specific. Downward adjustments may be more counterintuitive than upward adjustments, and cause adjustments to be particularly insufficient for low profit products.

In these experiments subjects made several rounds of decisions, received feedback after each round, and had the opportunity to learn from experience. Average order quantities, however, did not adjust toward the expected profit-maximizing quantity across rounds. This lack of learning is somewhat surprising. In these experiments subjects received quick feedback in a simple and stable environment. This also suggests that learning is even less likely in real world situations. In those cases managers face several impediments to learning from experience. Managers must operate in dynamic environments, and past lessons may provide little guidance for future decisions. Further, the quality of a newsvendor decision is often difficult to assess in hindsight: forgone sales are often difficult to observe and measure, unusual realizations of demand may skew evaluations of prior decisions (Hershey and Baron 1992), and disentangling current knowledge from what was known at the time of the decision is extremely difficult (Fischhoff 1982).

With these results in mind, we propose techniques to help firms manage the inventory decision process. The goal is to improve the inventory decisions of managers, but some of these techniques improve the quality of a decision by accounting for systematic bias, rather than actually correcting it.

One approach to improving inventory order decisions is to separate the forecasting task from the inventory decision task. While the forecasting task typically requires managerial judgment, the task of converting a forecast into an order quantity can be automated. A firm may reduce decision bias by asking managers to generate forecasts that are then automatically converted into order quantities. For example, a firm could determine the appropriate critical fractile, $(p - c)/(p - s)$, and elicit a forecast for the corresponding level of demand. By ordering this amount, the firm may avoid bias in the conversion task. This approach is similar to the two-stage decision process employed by L.L. Bean, Inc.® (see Schleifer 1993).

Automating the decision task is relatively simple, and should significantly improve decision quality. However, managers may be uncomfortable relinquishing control over the decision process. In addition, if managers recognize that their forecasts will automatically be converted to order quantities, they may "game" the system by altering their forecasts to target a particular order quantity. Ultimately, this may reintroduce the too low/too high pattern into order decisions.

Changing managerial incentives is a second approach to improving inventory order decisions. Results from this work suggest that managers are sensitive to profit margins, but tend to understock high profit products and overstock low profit products. To counteract these tendencies, a firm could implement corrective incentives such as penalties or bonuses. For example, for high-profit products a firm could impose a stockout penalty or a subsidy for leftover inventory. Analogously, for low-profit products a firm may lower orders by imposing a stockout bonus or an excess inventory penalty.

Determining the appropriate magnitude of these incentives, however, may be difficult. While these incentive schemes will shift inventory orders in the direction of the optimal order quantity, actual adjustments may be too small or too large. In a static, repeated environment, firms may collect information to determine the appropriate incentive level. Even individual differences, however, will influence the effectiveness of this approach.

A third approach involves training. Managers may be able to improve their inventory decisions with training and experience. While results from this work suggest that some types of training are ineffective, such as exposure to course material, other approaches may improve decision quality. Warning managers of the too low/too high pattern, providing them with practice and giving them feedback may improve decision quality. In particular, managers should receive profit and cumulative profit information for their own decisions as well as for alternative order quantities. Combined with bias training, this approach may improve decision quality. In a different domain, Friedman (1998) found that feedback and incentives substantially improved decision quality, though even very explicit feedback did not eliminate systematic errors. Note that practice and feedback will be most effective when managers make similar types of decisions in a repeated environment and when excess demand is easily observed.

4.1. Conclusion

The expected profit-maximizing solution for the newsvendor problem is well documented in the literature, but this is the first research to investigate newsvendor decision making. This research direction deserves greater attention since many managers use their judgment to modify recommendations made by models or make decisions that are completely unassisted by models. In particular, experimental methodology should be used to study other critical models in operations management. For example, workers are likely to change their behavior once firms implement new production systems (e.g., JIT, kanban, cellular manufacturing, and U-shaped production lines). It is important to confirm that workers change their behavior in a manner that is consistent with the objectives and assumptions of the new system. For example, Schultz et al. (1998) found that traditional assumptions do not adequately describe the actual behavior of production line workers in their experiments. Reengineering of work (Hammer 1990) presents another example in which the effectiveness of a new process design can depend on the behavioral responses of the workers involved with the process. In the queuing literature there has been some research that investigates the preferences of customers in a queuing system (Carmon et al. 1995), but no work on the actions of managers attempting to control or design a queuing system. A better understanding of actual behavior in real processes may lead to the discovery that traditional assumptions need modification, and that new techniques may be required to correctly optimize these systems.[7]

[7] The order of authorship was chosen randomly. The authors would like to thank seminar participants at the Wharton School, Colin Camerer, Karen Chinander, Kenneth Doer, and Chip Heath for helpful comments. The helpful comments of the Associate Editor and the referees are also graciously acknowledged. Support for this work was provided to Maurice Schweitzer through a McLamore Award in Business and the Social Sciences and a General Research Support Award from the University of Miami.

References

Arkes, H. R. 1996. The psychology of waste. *J. Behavioral Decision Making* **9** 213–224.

Bell, D. 1982. Regret in decision making under uncertainty. *Oper. Res.* **30** 961–981.

—— 1985. Disappointment in decision making under uncertainty. *Oper. Res.* **33** 1–27.

Beaty, S. 1996. Some Burger Kings fall short of free toy. *Wall Street Journal* November 25 B1.

Cachon, G., M. Schweitzer. 1998. Inventory forecasting and decision making with an unknown demand distribution. Working Paper, The Wharton School, University of Pennsylvania, Philadelphia, PA.

Carlson, J., T. O'Keefe. 1969. Buffer stocks and reaction coefficients: An experiment with decision making under risk. *Rev. Econ. Stud.* **36** 467–484.

Carmon, Z., Shanthikumar, J., T. Carmon. 1995. A psychological perspective on service segmentation models: The significance of accounting for consumers' perceptions of waiting and service. *Management Sci.* **41** 1806–1815.

Croson, R., K. Donohue. 1998. Institutional factors in supply chain management: An experimental study on reducing the bullwhip effect. Working Paper, The Wharton School, University of Pennsylvania, Philadelphia, PA.

Diehl, E., J. Sterman. 1995. Effects of feedback complexity on dynamic decision making. *Organ. Behavior and Human Decision Processes.* **62** 198–215.

Eeckhoudt, L., Gollier, C., H. Schlesinger. 1995. The risk-averse (and prudent) newsboy. *Management Sci.* **41** 786–794.

—— 1991. Increases in risk and deductible insurance. *J. Econom. Theory.* **55** 435–440.

Fischhoff, B. 1982. For those condemned to study the past: Reflections on historical judgment. D. Kahneman, P. Slovic, A. Tversky, eds. *Judgment under Uncertainty: Heuristics and Biases.* Cambridge University Press, New York, 335–351.

Fisher, M., A. Raman. 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* **44** 87–99.

Friedman, D. 1998. Monte Hall's three doors: Construction and deconstruction of choice anomaly. *Amer. Econom. Rev.* **88** 933–946.

Hammer, M. 1990. Reengineering work: Don't automate, obliterate. *Harvard Bus. Rev.* **68**(4) 104–112.

Hershey, J. C., J. Baron. 1992. Judgment by outcomes: When is it justified? *Organizational Behavior and Human Decision Processes.* **53** 89–93.

Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica.* **47** 263–291.

——, P. Slovic, A. Tversky. 1982. *Judgment under Uncertainty: Heuristics and Biases.* Cambridge University Press, New York.

Rosenfield, D. 1986. Optimal management of tax-sheltered employee reimbursement programs. *Interfaces* **16** 68–72.

Schleifer, A. Jr. 1993. L.L. Bean Inc: Item forecasting and inventory management. Harvard Business School Case 9-893-003. Cambridge, MA.

Schultz, K., D. Juran, J. Boudreau, J. McClain, L. J. Thomas. 1998. Modeling and worker motivation in JIT production systems. *Management Sci.* **44** 1595–1607.

Sterman, J. 1989. Modeling managerial behavior: Misperception of feedback in a dynamic decision making experiment. *Management Sci.* **35** 321–339.

Ziegler, B. 1994. IBM sells out new Aptiva pc; Shortage may cost millions in potential revenue. *Wall Street Journal.* October 7 B4.

—— 1995. IBM's O'Malley resigns post at pc division. *Wall Street Journal* February 7. B4.

Accepted by L. Robin Keller; received June 28, 1996. This paper was with the authors $15\frac{1}{2}$ *months for 4 revisions.*