Reference Dependence in Multilocation Newsvendor Models: A Structural Analysis

Teck-Hua Ho
Haas School of Business, University of California, Berkeley, Berkeley, California 94720, hoteck@haas.berkeley.edu

Noah Lim
Wisconsin School of Business, University of Wisconsin–Madison, Madison, Wisconsin 53706, nlim@bus.wisc.edu

Tony Haitao Cui
Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455, tcui@umn.edu

We propose a behavioral theory to predict actual ordering behavior in multilocation inventory systems. The theory rests on a well-known stylized fact of human behavior: people's preferences are reference dependent. We incorporate reference dependence into the newsvendor framework by assuming that there are psychological costs of leftovers and stockouts. We also hypothesize that the psychological aversion to leftovers is greater than the disutility for stockouts. We then experimentally test the proposed theory in both the centralized and decentralized inventory structures using subjects motivated by substantial financial incentives. Consistent with the proposed theory, actual orders exhibit the so-called "pull-to-center" bias and the degree of bias is greater in the high-profit margin than in the low-profit margin condition. These systematic biases are shown to eliminate the risk-pooling benefit when the demands across store locations are strongly correlated. Because the proposed model nests the standard inventory and ex post inventory error minimization theories as special cases, one can systematically evaluate the predictive power of each alternative using the generalized likelihood principle. We structurally estimate all three theories using the experimental data, and the estimation results strongly suggest that the proposed behavioral theory captures actual orders and profits better. We also conduct two experiments to validate the behavioral model by manipulating the relative salience of the psychological costs of leftovers versus that of stockouts to alleviate the pull-to-center bias.

Key words: multilocation inventory system; newsvendor; behavioral operations management; experimental economics

History: Received February 6, 2009; accepted June 12, 2010, by Martin Lariviere, operations management. Published online in *Articles in Advance* October 11, 2010.

1. Introduction

One of the most striking empirical findings in operations management research is that decision makers appear to have great difficulty arriving at the optimal order prescribed by standard inventory theory, even in the simplest setting of a single-store newsvendor. This finding was first reported by Schweitzer and Cachon (2000), who used laboratory economic experiments to study how subjects with managerial experience make ordering decisions for a single-store newsvendor. They documented a "pull-to-center" bias, that is, relative to the optimal order, subjects overorder when the ordering cost is high and underorder when the ordering cost is low. Bolton and Katok (2008) and Bostian et al. (2008) conducted follow-up experiments and showed that the pull-to-center bias is robust to changes in the features of the experimental design such as increasing the monetary payoffs, increasing the number of decision rounds, simplifying the decision space, and providing summary statistics of past demands.

If managers are in charge of multiple store locations, as is often the case in practice, then their ordering decisions are even more complex. In the multilocation setting, they not only have to take into account multiple store demands that may be correlated, but also consider the question of whether to centralize the inventory structure. Standard inventory theory prescribes a centralized inventory structure over a decentralized one as long as demands across the locations are not perfectly correlated. This is due to the benefit of risk pooling (Eppen 1979). However, the ability of a firm to fully realize this benefit has not been empirically verified. Moreover, the standard analysis that generates the risk-pooling benefit assumes that managers who make the ordering decisions are perfectly rational and will not exhibit systematic biases. As mentioned above, this perfect rationality assumption has been challenged by recent empirical findings that ordering decisions are biased in the setting of a single-store newsvendor problem. Given these findings, one would expect that orders in the multistore setting

would also not conform to the theoretical prediction. One potential implication is that firms may not realize the full benefits of risk pooling because of the biases in order quantities. In this paper, we test whether the same kind of pull-to-center bias exists in a multilocation newsvendor setting.

Schweitzer and Cachon (2000) showed that if a manager has a preference for minimizing ex post inventory error, then she will exhibit the pull-to-center bias. The authors however did not estimate their behavioral model using the experimental data. In this paper, we propose a new behavioral model that rests on a well-known empirical fact about human beings: people's preferences are reference dependent (Kahneman and Tversky 1979; Thaler 1985; Kőszegi and Rabin 2006; and Ho et al. 2006a, b). We incorporate reference dependence into the newsvendor framework by assuming that there are psychological costs of leftovers and stockouts. We also hypothesize that the psychological aversion to leftovers is greater than the disutility for stockouts. Our model nests the perfect rationality hypothesis and Schweitzer and Cachon's (2000) model as special cases. These direct relationships allow us to structurally estimate all three models simultaneously using our experimental data. In this way, we provide a deeper understanding of the behavioral underpinnings for the pull-to-center bias.

This paper contributes to the emerging field of behavioral operations management in four ways:

1. We develop a new behavioral theory of a multilocation inventory system. Our model assumes that in addition to the standard pecuniary costs of having leftovers and not meeting demand, decision makers also experience psychological disutility from having leftovers and stockouts. We allow these two components of disutility to be weighted differently because prior research suggests that the psychological disutility from losses that are actually incurred (for the case of leftovers) are stronger than forgone losses (for the case of stockouts). We prove that when decision-makers exhibit reference dependence, the model predicts a pull-to-center bias in both the centralized and decentralized inventory structures. We also show that when the psychological cost of a leftover is higher than that of a stockout, the pull-to-center bias is stronger in the high-profit margin relative to the low-profit margin condition. We derive closed-form solutions for the order quantities given the behavioral biases.

2. We design and conduct the first experiment to examine orders and profit outcomes of the multilocation newsvendor across the centralized and decentralized inventory structures. We also vary the levels of the product profit margin (low or high, which we term the low-profit and high-profit conditions). Thus, our main experiment consists of four separate treatments. Within each treatment, we also test the robustness of the theory predictions using two different scenarios of demand correlations among store locations—one where demands are uncorrelated versus one where they exhibit strong positive correlations (with respective correlation coefficients of 0 and 0.8). The data show that the pattern of orders across all treatments is consistent with the predictions of the generalized newsvendor model, with pull-to-center biases in all four treatments and a stronger bias in the high-profit conditions. Consequently, the actual average profits are lower than predicted by the standard inventory model. We also find that when the demands across stores are uncorrelated, actual profits under centralization are indeed higher relative to decentralization, so the risk-pooling benefit survives the pull-to-center in orders. However, when the demands are strongly correlated, we did not find statistical support for the risk-pooling benefit.

3. We structurally estimate the proposed behavioral model using the experimental data and show that the model tracks actual orders more closely than the standard inventory model and Schweitzer and Cachon's (2000) model. The parameter estimates indicate that the psychological cost of a leftover is 1.53 times greater than the psychological cost of a stockout, which also explains the stronger pull-to-center bias in the high-profit conditions. The approach of structural estimation is rare in the behavioral operations management literature and our paper is one of the first to estimate the behavioral parameters of a model with psychological biases.[1]

4. We validate the new behavioral theory using two follow-up experiments. The key idea behind these experiments is to attempt to manipulate the relative salience of the psychological cost of a leftover versus the psychological cost of a stockout, such that the pull-to-center bias is predicted to be reduced if the psychological drivers of the behavioral model are indeed operational. Which one of the two psychological costs we make salient to subjects depends on whether it is the low-profit or high-profit condition (so that we know the direction in which to move the ordering decisions). In the first experiment, we asked subjects to compute and write down the costs of leftovers (unmet demands) in the low-profit (high-profit) condition in every round if these costs are incurred in that round. In the second experiment, we introduced a small monetary penalty (bonus) for having leftovers (for meeting all customer demand) in the low-profit (high-profit) conditions. The results of both

---

[1] The other exceptions are Su (2008), who estimated the error parameter in a model where decision makers are allowed to make computational errors and Bostian et al. (2008), who estimated the parameters of a learning model (Camerer and Ho 1999) that tracks how people adjust orders over multiple decision rounds.

experiments show that the pull-to-center bias is alleviated in aggregate and for most individual subjects. These validation tests further support the reference-dependence theory.

The rest of this paper is organized as follows. Section 2 introduces the reference-dependence model for the multilocation newsvendor. We contrast the predictions of this model with those of the standard inventory model using three propositions. Section 3 describes the design and results of the main experiment, the structural estimation of the reference-dependence model, and shows that the generalized behavioral model provides a superior explanation of the actual pattern of orders and profits. Section 4 presents the two experiments designed to validate our behavioral model. We offer some concluding remarks in §5.

2. Reference Dependence in the Multilocation Newsvendor

A newsvendor is in charge of n retail stores that sell an identical product. Each store has a demand (in units), denoted D_i for store $i = 1, \ldots, n$. The demands across the stores are identically distributed following a normal distribution $F(\cdot) \sim N(\mu, \sigma^2)$, and the correlation between any two stores is given by ρ $(-1/(n-1) \le \rho \le 1)$. The newsvendor's decision is to order the products before the selling cycle begins in order to maximize expected profit. The unit retail price is p, the unit cost is c, and s is the unit salvage value of unsold units. The price and unit cost are exogenously determined and independent of the sales volume.

If this multilocation newsvendor faces a decentralized inventory structure, then orders are specific to each store, so it orders q_i^d for store $i = 1, \ldots, n$. Because the retail price p and unit cost c are constant across all stores, the newsvendor will order the same quantity $q_i^d = q^d$ for each of the n stores. The aggregate ordering quantity across stores is $Q^d = n \cdot q^d$. The newsvendor's total profit under decentralization is $\Pi^d = n \cdot [p \cdot \sum_{i=1}^n \min[q^d, D_i] - c \cdot q^d]$. If the inventory is centralized however, that is, the product units that are ordered can be freely allocated across the n stores after the demand realizations, the newsvendor determines an aggregate ordering quantity Q^c. Because demands across stores are identically distributed with a correlation ρ, if $F(\cdot)$ has a mean of μ and a variance of σ^2, then the total demand distribution under centralization, $G(\cdot)$, has a mean of $n \cdot \mu$ and a variance of $[n + n(n-1)\rho] \cdot \sigma^2$. The newsvendor's total profit in this case is given by $\Pi^c = p \cdot \min[Q^c, \sum_{i=1}^n D_i] - c \cdot Q^c$.

Under both centralization and decentralization, the goal of the newsvendor is to maximize its expected profit $E\Pi(\cdot)$ given the corresponding demand distribution, which is $F(\cdot)$ under decentralization and $G(\cdot)$ under centralization. It is customary to attack each of the profit maximization problems by solving an equivalent optimization problem that minimizes the corresponding expected costs. In the newsvendor context, there are two types of costs—the costs of overordering (from having leftovers) and the costs of underordering (from having stockouts). Under standard inventory theory, the manager arrives at the optimal ordering decision through minimizing the sum of these two types of costs.

However, there is a large stream of research that has shown that people make decisions not only based on their final pecuniary outcomes, but also based on changes in outcomes relative to a reference point (Kahneman and Tversky 1979). Reference-dependent preferences have been used to explain a broad range of behavioral "anomalies" such as why cab drivers do not work longer hours on days when customers are plentiful (they have a target income level and stop working beyond that target), why contestants increase effort in contests with more winners than losers (losing hurts more when contestants are now expected to win), and why sellers overprice their houses in a real estate bust (they try to avoid making a loss relative to their purchase price).[2] In the newsvendor model, the manager knows that he will be confronted by the demand realizations after the ordering decision, and that the losses from overordering or underordering will be computed based on the actual demands. Hence, these realized demands serve as natural focal points by which the manager evaluates the effectiveness of his ordering decision. Besides the pecuniary costs of having leftovers or stockouts, the manager may also experience additional psychological disutility because the firm's losses resulted from his ordering decision. Moreover, the degree of psychological pain may differ depending on whether the losses are actual losses (from having leftovers) or forgone losses (from having stockouts). Prior research has shown that decision makers underweight or even neglect forgone payoffs, both in individual decision-making tasks (Thaler 1980, Frederick et al. 2009) and in strategic situations (Camerer and Ho 1999). Applying this finding to the newsvendor model, one could posit that the psychological pain from having a leftover may be greater than that from having a stockout.

In this paper, we capture reference dependence in the newsvendor setting by applying the modeling framework of Thaler (1985) and Kőszegi and Rabin (2006) and specifying preferences as the sum of two components: an intrinsic utility component associated with the actual final outcome, and a change utility

---

[2] See Camerer et al. (1997), Lim (2010), and Genesove and Mayer (2001), respectively, for details of these examples. See also Ho et al. (2006a) for other applications of reference dependence.

component associated with gains and losses relative to a reference point.[3] Specifically, we assume that each of the two types of costs faced by the decision maker in the newsvendor problem consists of two components—an actual cost component and a psychological cost component that depends on a reference point. For the psychological cost component of utility, we use the realized demands as the reference points because, as mentioned above, they are the most critical and salient pieces of information the decision maker expects to receive after making the ordering decision. In other words, we assume that there will be no psychological disutility only when orders are equal to the realized demands, that is, when there are zero overage and underage costs.[4] Formally, the expected utility $EU(Q_r)$ as a function of the ordering quantity Q_r ($Q_r = Q_r^d/n = q_r^d$ under decentralization or Q_r^c under centralization)[5] for a decision maker with reference-dependent preferences facing a demand distribution with probability density function (p.d.f.) $h(x)$ is given by

$$EU(Q_r) = \int_0^{Q_r} [px + s(Q_r - x) - \delta_o(Q_r - x)] \cdot h(x)\, dx$$
$$+ \int_{Q_r}^{\infty} [pQ_r - \delta_u \cdot (x - Q_r)] \cdot h(x)\, dx - cQ_r, \quad (1)$$

where $\delta_o \geq 0$ is the psychological per-unit cost of overordering for $Q_r > D$, $\delta_u \geq 0$ is the psychological per-unit cost of underordering for $Q_r \leq D$, and $H(\cdot)$ is the focal distribution of demand that the newsvendor uses to make orders, that is, $H(\cdot) = G(\cdot)$ under centralization and $H(\cdot) = F(\cdot)$ under decentralization. If the decision maker underweights forgone payoffs as suggested by previous research, we have $\delta_o > \delta_u$. Note that when $\delta_o = \delta_u = 0$, there is no reference dependence and Equation (1) reduces to the standard inventory model.

With the usual notations, we have underage cost $c_u = (p - c)$ and overage cost $c_o = (c - s)$. The normal distribution is symmetric, so that $F(\mu) = 1/2$ in the decentralized scenario and $G(n \cdot \mu) = 1/2$ under centralization, where μ is the expected demand at each store. The decision maker maximizes expected utility, and the optimal order quantity with reference-dependent preferences can be solved through the

---

[3] Thaler (1985) proposed in the context of consumer choice that a consumer's overall utility from a purchase is composed of "acquisition utility," which is the consumer surplus in standard economic models, and also "transaction utility," which is the perceived value of the deal relative to some reference price.

[4] Ho et al. (2007) show that using the best-payoff scenario ex post as the reference point by which decision makers evaluate their outcomes predicts learning behavior in strategic situations well.

[5] Here $EU(Q_r)$ is equal to the sum of $E\Pi(Q_r)$ and the psychological cost of overordering for $Q_r > D$ or underordering for $Q_r < D$, where $E\Pi(Q_r)$ is the per-store expected profit under decentralization and the total expected profit under centralization.

first-order condition (FOC) of Equation (1) and is given by

$$H(Q_r^*) = P(X \leq Q_r^*) = \frac{c_u + \delta_u}{c_u + \delta_u + c_o + \delta_o}$$
$$= \frac{p - c + \delta_u}{p - s + \delta_u + \delta_o} = cr, \quad (2)$$

where cr is the so-called *critical ratio*.

Note that if we restrict the behavioral model by assuming that when people make actual ordering decisions, they care only about minimizing actual costs and will not be influenced by reference dependence (i.e., $\delta_o = \delta_u = 0$), we recover the standard inventory model with the optimal order quantity Q^* given by the familiar expression of

$$H(Q^*) = P(X \leq Q^*) = \frac{c_u}{c_u + c_o} = \frac{p - c}{p - s}. \quad (3)$$

We denote Q^{c*} as the optimal total order under centralization and q^{d*} as the optimal order per store under decentralization in the standard inventory model.

2.1. Optimal Orders

Under decentralization, the optimal order for each store is

$$q_r^{d*} = \mu + Z_{cr, r} \cdot \sigma, \quad (4)$$

where $Z_{cr, r}$ is the inverse of the cumulative normal density evaluated at the critical ratio $(p - c + \delta_u)/(p - s + \delta_u + \delta_o)$ of the behavioral model in Equation (2). The total order for the n stores is simply equal to

$$Q_r^{d*} = n \cdot q_r^{d*} = n \cdot \mu + n \cdot Z_{cr, r} \cdot \sigma. \quad (5)$$

Under centralization, orders are given by

$$Q_r^{c*} = \mu^c + Z_{cr, r} \cdot \sigma^c$$
$$= n \cdot \mu + Z_{cr, r} \cdot \sigma \sqrt{n + n(n - 1)\rho}, \quad (6)$$

where again $Z_{cr, r}$ is the inverse of the cumulative normal density evaluated at the critical ratio $(p - c + \delta_u)/(p - s + \delta_u + \delta_o)$ of the behavioral model in Equation (2). Furthermore, we need to have $n + n(n - 1)\rho \geq 0$ in order to have σ^c as a real number, which leads to $-(1/(n - 1)) \leq \rho \leq 1$.

Proposition 1 below shows that incorporating psychological costs due to reference dependence can account for pull-to-center effects in orders. Also, following Schweitzer and Cachon (2000), we define the profit condition faced by the newsvendor to be a *high-profit* condition when $cr \geq 1/2$ and a *low-profit* condition otherwise under the assumption of $\delta_o = \delta_u = 0$. We assume constant price p and salvage value s, and we denote the cost in the high-profit condition as c_{hp} and the cost in the low-profit condition as c_{lp} when we

compare across profit conditions. We thus have $c_{lp} > c_{hp}$ because the critical ratio in the high-profit condition is larger than that in the low-profit condition.

Denote $\Omega_{lp}^c = Q_r^{c*} - Q^{c*}$ and $\Omega_{lp}^d = q_r^{d*} - q^{d*}$ to measure the degree of pull-to-center in the low-profit condition, and denote $\Omega_{hp}^c = Q^{c*} - Q_r^{c*}$ and $\Omega_{hp}^d = q^{d*} - q_r^{d*}$ to measure the degree of pull-to-center in the high-profit condition. The degrees of pull-to-center are therefore given by

$$
\begin{aligned}
\Omega_{lp}^c &= (Z_{cr,r}^{lp} - Z_{cr}^{lp}) \cdot \sigma \sqrt{n + n(n-1)\rho}, \\
\Omega_{lp}^d &= (Z_{cr,r}^{lp} - Z_{cr}^{lp}) \cdot \sigma, \\
\Omega_{hp}^c &= (Z_{cr}^{hp} - Z_{cr,r}^{hp}) \cdot \sigma \sqrt{n + n(n-1)\rho}, \\
\Omega_{hp}^d &= (Z_{cr}^{hp} - Z_{cr,r}^{hp}) \cdot \sigma,
\end{aligned} \tag{7}
$$

where $Z_{cr,r}^{lp}$, Z_{cr}^{lp}, $Z_{cr,r}^{hp}$, and Z_{cr}^{hp} are the inverses of the cumulative normal density evaluated at the corresponding critical ratios under the behavioral and standard inventory models.

PROPOSITION 1. *In the low-profit condition, orders exhibit pull-to-center, i.e., $\Omega_{lp}^c > 0$ and $\Omega_{lp}^d > 0$, as long as*

$$
\frac{\delta_u}{\delta_o} > \frac{p - c_{lp}}{c_{lp} + s}, \tag{8}
$$

and in the high-profit condition, orders exhibit pull-to-center, i.e., $\Omega_{hp}^c > 0$ and $\Omega_{hp}^d > 0$, as long as

$$
\frac{\delta_u}{\delta_o} < \frac{p - c_{hp}}{c_{hp} + s}, \tag{9}
$$

for any $\delta_u \geq 0$, $\delta_o \geq 0$, and $c_{lp} > c_{hp}$.

PROOF. See Appendix A.[6]

Next, we show that if the psychological per-unit cost of overordering is larger than that of underordering, that is, $\delta_o > \delta_u$, there will be an *asymmetric* pull-to-center effect when we compare a low-profit condition with its *symmetric* high-profit condition. We define a low-profit condition and a high-profit condition to be symmetric if the optimal orders in both conditions have the same distance to the expected demand under the assumption of $\delta_o = \delta_u = 0$.

PROPOSITION 2. *The pull-to-center effect is stronger in the high-profit condition relative to its symmetric low-profit condition, $\Omega_{hp}^c > \Omega_{lp}^c$ and $\Omega_{hp}^d > \Omega_{lp}^d$, if and only if $\delta_o > \delta_u > 0$.*

The above propositions show how orders will be biased if decision makers indeed have reference-dependent preferences. The following proposition characterizes the implications on expected profits under reference dependence.

PROPOSITION 3. *The ordering quantity with reference-dependence biases leads to a lower profit than the optimal order in the standard model, and the profit difference is given by*

$$
\begin{aligned}
&E\Pi(Q^{c*}) - E\Pi(Q_r^{c*}) \\
&= (c - s)(Q_r^{c*} - Q^{c*}) \\
&\quad + (p - s)\sigma_H \left[R\left(\frac{Q_r^{c*} - \mu_H}{\sigma_H}\right) - R\left(\frac{Q^{c*} - \mu_H}{\sigma_H}\right) \right]
\end{aligned} \tag{10}
$$

in the centralized inventory structure and

$$
\begin{aligned}
&E\Pi(q^{d*}) - E\Pi(q_r^{d*}) \\
&= n(c - s)(q_r^{d*} - q^{d*}) \\
&\quad + n(p - s)\sigma_H \left[R\left(\frac{q_r^{d*} - \mu_H}{\sigma_H}\right) - R\left(\frac{q^{d*} - \mu_H}{\sigma_H}\right) \right]
\end{aligned} \tag{11}
$$

in the decentralized inventory structure, where μ_H and σ_H^2 are the mean and variance of the focal demand with a distribution $H(\cdot)$ and $R(\tau) = \int_\tau^\infty (y - \tau)(1/\sqrt{2\pi})e^{-(y^2/2)} \, dy$.

It is important to mention at this point that one may generate a similar pull-to-center result in orders using alternative explanations. The most well known of these is the model with preferences for minimizing ex post inventory error by Schweitzer and Cachon (2000). We note that interestingly, whereas our model is based on different psychological underpinnings, it can be shown to be a mathematical generalization of their model for the special case of $\delta_o = \delta_u = \delta$. To see this, let $\delta_u = \delta_o = \delta$ and rewrite Equation (1) as

$$
\begin{aligned}
EU(Q_r) &= \int_0^{Q_r} [px + s(Q_r - x)] \cdot h(x) \, dx \\
&\quad + \int_{Q_r}^\infty pQ_r \cdot h(x) \, dx - cQ_r \\
&\quad - \int_0^{Q_r} \delta \cdot (Q_r - x)h(x) \, dx \\
&\quad - \int_{Q_r}^\infty \delta \cdot (x - Q_r)h(x) \, dx \\
&= \int_0^{Q_r} [px + s(Q_r - x)] \cdot h(x) \, dx \\
&\quad + \int_{Q_r}^\infty pQ_r \cdot h(x) \, dx - cQ_r \\
&\quad - \int_0^\infty \delta \cdot |Q_r - x|h(x) \, dx,
\end{aligned} \tag{12}
$$

where $-\int_0^\infty \delta \cdot |Q_r - x|h(x) \, dx$ represents the disutility of the ex post inventory error with a per-unit psychological cost of δ. Note that an empirical implication of the model of Schweitzer and Cachon (2000) is that Proposition 2 will not hold, that is, the pull-to-center effect will be equally strong across symmetric low-profit and high-profit conditions.

---

[6] All the proofs that are not given in the main text of the paper can be found in Appendix A.

To determine the relative efficacy of the predictions of the reference-dependence model versus its nested models, that is, the model of Schweitzer and Cachon (2000) and the standard inventory model, we conduct an incentive-aligned experiment to study how people make ordering decisions in a multilocation newsvendor setting. Based on our discussion of Propositions 1–3, the objectives of the experiment are to assess the following questions: First, is there a pull-to-center effect in the actual ordering decisions across both the centralized and decentralized inventory structures and also across the low-profit and high-profit conditions, consistent with the prediction of Proposition 1? Second, is there a stronger pull-to-center effect in the high-profit condition relative to the low-profit condition, consistent with the prediction of Proposition 2 when $\delta_o > \delta_u$? Third, are the actual profits lower than the levels predicted by the standard inventory model as predicted by Proposition 3? We describe the details of the experimental design and the results in the next section.

3. Experimental Test

3.1. Design and Procedure
To test if the predictions of the reference-dependence model are valid, we employ a 2×2 between-subjects experimental design that varies the type of inventory structure ("centralized" or "decentralized") and the product profit condition ("low" or "high," through varying the unit cost c). We varied the inventory structure because we wanted to check if our model is robust across the two types of inventory structure that a multilocation newsvendor may face—a centralized or a decentralized one. We varied the profit condition because we wanted to test if the psychological costs of leftovers and stockouts are different through comparing the strength of the pull-to-center effect (if any) across the profit conditions. As an additional robustness check, we also implemented two different scenarios of demand correlations between stores ($\rho = 0$ and $\rho = 0.8$) within each of these four treatments.

Subjects were undergraduate business majors at a public research university in the United States and all of them had completed at least one statistics course at the undergraduate level. A total of $S = 160$ subjects were recruited, with 40 subjects assigned to each of the four treatment cells. Each subject made ordering decisions for 40 rounds. Two experimental sessions were conducted for each treatment, with each session corresponding to a different demand correlation scenario. Each session has 20 subjects and lasted approximately 75 minutes. Subjects received course credit for arriving on time and earned experimental points that were converted into cash at the end of the session. The average dollar earnings were $16.80, and

the minimum and maximum earnings were $2 and $24, respectively.

Once subjects entered the room, each of them was assigned an identification number (that runs from 1 to 20 in each session) and seated at a computer terminal. The instructions were then read aloud by the experimenter. Each subject was told that she was a manager in charge of four retail stores—A, B, C, and D. The stores generated profits by selling products to consumers, but the products must first be ordered from a supplier and sent to the stores before the selling season began. Subjects were told that their decision in each round was to decide on how many units to order (this was called *Order*). In the centralized treatments, each subject made one order for all stores. To equalize the number of decisions that subjects made across the two different inventory structures, each subject also made only one order in the decentralized treatments—that order quantity was applied to each of the four stores. The retail price p for each unit of the product sold was 10 points and the cost of each unit ordered from the supplier was set at either $c_{lp} = 8$ or $c_{hp} = 2$ points, corresponding to the low-profit and high-profit conditions, respectively. Note that these parameter values yield a symmetric low-profit and high-profit design. To simplify the decision arithmetic, we set the salvage value s to be 0.

Next, subjects were told that the number of units that would be sold depended on the consumer demand for each store, which was uncertain at the time of the ordering decision and would be generated by the computer only after the order was placed. The consumer demand for each store had a distribution that was approximately normal,[7] with a mean of $\mu = 1,000$ units and a standard deviation of $\sigma = 400$ units. Subjects were also told that the median and modal demand was 1,000 and were provided a table that listed the chance (in percentages) that demand would fall within a particular range, with the range varying in increments of 200 units, from 0 up to 1,800 units. For demand ranges above 1,800, we simply included a row in the table that indicated that the percentage chance that demand for each store would range from "1,801 units and above" was 2.3%. From this table, subjects were also able to discern the general shape and the symmetry of the distribution—for example, they were able to tell that the chances

---

[7] We used the term "approximately normal" because there was a 0.625% chance that the demand for a store would be less than 0 units. When generating the demand draws, we resampled the distribution whenever we encountered a demand that was less than 0. Because demand was always constrained to be positive and, as we described the distribution as "approximately normal," preferring it to the more technical term of "truncated normal distribution" to avoid confusing subjects who might not be familiar with that term. All subjects were familiar with the normal distribution.

that demand for a store would fall within 601 to 800 and 1,201 to 1,400 were equal at 15.0%. In addition, all subjects were allowed to view an identical sample of 30 demand observations. In the experiment, the demand values generated varied across the twenty subjects within each session. However, subjects with the same identification numbers across the four treatments for the same given level of ρ received the same set of demand values (naturally, the set of demand values were different for the different scenarios of $\rho = 0$ and $\rho = 0.8$). This design feature is particularly useful because it will allow us to compare profits across treatments while controlling for the impact of demand realizations on actual profits.

For the scenario where $\rho = 0$, we simply told subjects in the instructions that the demands of stores A, B, C, and D in each round were independently generated, clarifying that the demand value of one store would not affect the demand value of another store. In the scenarios where $\rho = 0.8$, we told subjects that the demands are positively correlated with a correlation coefficient of 0.8, and also explained what the correlation coefficients of 0 and 1 meant. To further illustrate the concept of correlation, we also provided subjects with scatter plots of the demands between two stores for these three values of ρ. Finally, in all the treatments, subjects were also told that for each store, the past values of demand are independent of future demand values, that is, there is no correlation in demands across decision rounds.

In the decentralized treatments, subjects were told that for each store, if the store's demand was less than *Order*, the number of units sold would be the former and the leftovers would not have any salvage value. Conversely, if the store's demand was greater than *Order*, the number of units sold would be the latter, so that there will be unmet demand. The point earnings in both these cases were carefully explained to the subjects. For example, for $c_{lp} = 8$, if *Order* was greater than the store's demand, the point earnings for that store would be equal to 10*(*Demand for that store*) − 8*Order*. Otherwise, the point earnings would be 10*Order* − 8*Order*. In each round, subjects were paid based on the sum of point earnings across the four stores. The information communicated in the centralized treatments was similar, except that *Order* was evaluated against the *Total Consumer Demand*, which was defined as the sum of the demands of the four stores. This was also the key distinction between the two inventory structures in our experimental instructions. For example, for $c_{hp} = 2$ under centralization, the point earnings when *Order* was less than the *Total Consumer Demand* was 10*Order* − 2*Order*.[8] The detailed instructions for the centralized

low-profit treatment for the $\rho = 0$ scenario can be found in Appendix B.

We faced two design challenges in converting the point earnings into cash. First, subjects should receive similar cash payments in each treatment so that stake size would not pose as a possible confound for any differences in decision performance across treatments.[9] Second, the decision task must be payoff dominant so that any deviation in the ordering quantity from optimality involved a substantial drop in profit (Smith 1982). To address the first challenge, we set a benchmark cash earnings level so that across all the four treatments, subjects would earn on average 40¢ in each round and receive an expected total sum of $20 (inclusive of a $4 start-up fee) for the 40 rounds if they order at the levels predicted by the standard inventory model. To address the second challenge, we performed an affine transformation of the cash conversion functions (which would not affect the optimal ordering decisions) so that subjects were required to exceed certain target point earnings in order to make a positive cash earning, and they made a negative cash earning if they failed to do so.

After the instructions were read, the experiment was implemented using a computer program specially written for the experiment. To familiarize them with the program and experimental procedure, subjects went through four practice rounds that carried no monetary consequences before the 40 decision rounds began. All subjects were given a start-up fee of $4. In every round, the computer displayed the round number, retail price, order cost, and the minimum points needed to achieve positive cash earnings on one side of the screen. Once subjects entered their ordering decisions, the computer displayed their ordering decisions, the demands for each of the four stores (and the *Total Consumer Demand* in the centralized treatments), the number of units sold for each store (with centralization, only the total number of units sold was displayed), the point and cash earnings for that round, and the cumulative cash earnings inclusive of the start-up fee. In addition, subjects were able to view all the decision and outcome histories of previous rounds. Hence, subjects did not have to compute their point and cash earnings after every round as these were performed by the computer. We computerized the earnings calculations because we wanted to eliminate computational errors (which

---

[8] In practice, the earnings of the managers who make the order decisions may not be directly aligned with firm profits. The reward

scheme in our experiment avoids this potential misalignment between individual and firm payoffs because we pay subjects the profits of the firm.

[9] In a comprehensive review of the effects of financial incentives, Camerer and Hogarth (1999) show that stake size often does not change the main pattern of results in individual decision tasks. Despite this, we choose to be cautious and control for stake size explicitly.

Table 1 Predicted Orders of the Standard Inventory Model and Actual Average (per Store)

Treatments	Scenario $\rho = 0$			Scenario $\rho = 0.8$		
	Theory	Actual	t-stat.[a]	Theory	Actual	t-stat.[a]
Centralized						
Low-profit (cost = 8)	832	941	5.4*	690	834	20.6*
High-profit (cost = 2)	1,168	992	−11.1*	1,310	1,032	−30.3*
Decentralized						
Low-profit (cost = 8)	663	813	4.9*	663	805	18.3*
High-profit (cost = 2)	1,337	1,057	−8.0*	1,337	1,098	−15.8*

[a]The t-statistics are the results of the tests of average orders against the predictions of the standard inventory model.

*Indicates significance at the 5% level.

Table 2 Predicted Expected Profits of the Standard Inventory Model and Actual Average Profits (Point Earnings)

Treatments	Scenario $\rho = 0$			Scenario $\rho = 0.8$		
	Theory	Actual	t-stat.[a]	Theory	Actual	t-stat.[a]
Centralized						
Low-profit (cost = 8)	5,760	4,633	−7.3*	3,870	3,606	−1.2
High-profit (cost = 2)	29,760	28,082	−5.2*	27,870	26,578	−3.8*
Decentralized						
Low-profit (cost = 8)	3,521	2,601	−6.0*	3,521	3,250	−1.2
High-profit (cost = 2)	27,521	25,265	−6.1*	27,521	26,190	−3.5*

Note. In the experiments, we use different point-earnings-to-cash-conversion ratios across the low-profit and high-profit treatments so that subjects will receive the same expected monetary payoff across the treatments if they order the quantities prescribed by the standard inventory model.

[a]The t-statistics are the results of the tests of average profits against the predictions of the standard inventory model.

*Indicates significance at the 5% level.

might also lead subjects to make suboptimal ordering decisions) and more importantly, allow the subjects to concentrate solely on the ordering task so that the predictions of the standard inventory model would have the greatest chance of succeeding. Once the subjects completed the 40 rounds, they were paid their cash earnings privately and directed to leave the room.

Tables 1 and 2 show, respectively, the point predictions for the order per store and the expected profits for all the four treatments under the standard inventory model where $\delta_o = \delta_u = 0$. Note that under decentralization, orders are not affected by the different ρ scenarios. We observe that under the standard inventory model, the optimal order on a per store basis under centralization is higher than under decentralization in the low-profit condition. In the $\rho = 0$ scenario, $q_{lp, \rho=0}^{c*} = 832$, which is higher than $q_{lp, \rho=0}^{d*} = 663$. Similarly, when $\rho = 0.8$, the centralized order is 690, compared to 663 under decentralization. In the high-profit condition ($c_{hp} = 2$), the pattern of orders is reversed. When $\rho = 0$, the optimal orders per store with centralization is $q_{hp, \rho=0}^{c*} = 1,168$, which is lower than $q_{hp, \rho=0}^{d*} = 1,337$ under decentralization. This is true also when $\rho = 0.8$: we have $q_{hp, \rho=0.8}^{c*} = 1,310$, which is lower than $q_{hp, \rho=0.8}^{d*} = 1,337$.

From Table 2, notice that the standard inventory model predicts expected profits to be higher under centralization across both profit conditions. This is the risk-pooling benefit of centralization mentioned at the beginning of the paper. In the ρ scenario, expected profits across the four stores under centralization are 2,239 points higher relative to decentralization across both profit conditions. In contrast, the difference in expected profits shrinks to 349 points when $\rho = 0.8$.

3.2. Results

The average order per store and the total profits (for all four stores) for the four treatments are reported

in Tables 1 and 2, respectively. Table 1 shows that there is a marked difference between actual average orders and the quantities predicted by the standard inventory model in all treatments. Orders in the two low-profit ($c_{lp} = 8$) treatments appear higher than predicted, whereas they are lower than predicted in the two high-profit ($c_{hp} = 2$) treatments. In other words, there is a pull-to-center pattern as predicted by the reference-dependence model. The average profits in Table 2 are directionally lower than the predictions of the standard inventory model in all four treatments. In the following paragraphs, we conduct formal statistical tests and summarize the main empirical findings. Because each subject made multiple decisions, we account for within-subject correlation by clustering the standard errors at the subject level in all the analyses.

RESULT 1. *Average orders in all four treatments differ significantly from the predictions of the standard inventory model and exhibit pull-to-center patterns.*

We begin by comparing the average orders across all rounds against the point predictions of the standard inventory model using one-sample t-tests. The results are reported in Table 1. First, average orders in all four treatments are significantly different from their predicted levels (all p-values 0.000). Second, there is a pull-to-center pattern in all treatments. That is, in the low-profit condition across inventory structures and the two scenarios of ρ, average orders are higher than predicted by the standard inventory model (all p-values 0.000). Correspondingly, in the high-profit condition, average orders across inventory structures and levels of ρ are lower than predicted (all p-values 0.000). Overall, the pattern of actual orders

Table 3   Results of OLS Regressions on the Degree of Pull-to-Center Across Profit Conditions

Treatment	Variable	Coefficient	t-stat.
Centralized $\rho = 0$ scenario	Constant (Base = Low-profit)	0.65 (0.12)[a]	5.35* (0.000)[b]
	High-profit	0.40 (0.15)	2.61* (0.013)
Decentralized $\rho = 0$ scenario	Constant (Base = Low-profit)	0.45 (0.93)	4.82* (0.000)
	High-profit	0.39 (0.14)	2.76* (0.009)
Centralized $\rho = 0.8$ scenario	Constant (Base = Low-profit)	0.46 (0.02)	20.42* (0.000)
	High-profit	0.43 (0.04)	11.55* (0.000)
Decentralized $\rho = 0.8$ scenario	Constant (Base = Low-profit)	0.42 (0.02)	18.19* (0.000)
	High-profit	0.29 (0.05)	5.67* (0.000)

Note. In each of the four regressions, there are 1,600 observations and 40 subject clusters.

[a]The numbers in parentheses below the coefficient estimates are the robust standard errors.

[b]The numbers in parentheses below the t-statistics are the p-values.

*Indicates significance at the 5% level.

is consistent with the prediction of the reference-dependence model as described in Proposition 1.

Result 2. *The pull-to-center effect is stronger in the high-profit condition, consistent with the prediction of the reference-dependence model when $\delta_o > \delta_u$.*

Proposition 2 in this paper predicts that if $\delta_o > \delta_u$, the pull-to-center effect will be stronger in the high-profit condition relative to its symmetric low-profit condition. To test for differences in the degree of pull-to-center across profit conditions (which are symmetric in our experiment), we construct a measure of the pull-to-center effect for the experimental data in the following way. For the high-profit treatments, we define $(q_{hp}^* - q_{hp,st})/(q_{hp}^* - 1{,}000)$ as a percentage measure of the degree of pull-to-center, where $q_{hp,st}$ is the per-store order of subject s in round t, and q_{hp}^* is the optimal per-store order under the standard inventory model for that high-profit treatment. For the low-profit condition, we define $(q_{lp,st} - q_{lp}^*)/(1{,}000 - q_{lp}^*)$ as the percentage measure of the degree of pull-to-center, where q_{lp}^* is the optimal order per store for that low-profit treatment. With this measure, a value of 0 indicates no pull-to-center and higher values represent stronger pull-to-center effects. Next, we compare the relative degree of the pull-to-center between the low-profit and high-profit conditions within each inventory structure and for the two different scenarios of ρ via four separate dummy-variable ordinary least squares (OLS) regressions (using the high-profit condition as the treatment dummy). The results in Table 3 show that the high-profit condition exhibits a stronger degree of pull-to-center in all cases, which supports the prediction of Proposition 2.

It is important to note that the order decisions summarized in the above two results are remarkably consistent at the individual subject level. Table 4 displays

Table 4   Average Orders (per Store) by Subjects Across 40 Rounds

	$\rho = 0$					$\rho = 0.8$			
	Centralized		Decentralized			Centralized		Decentralized	
Subject ID[a]	Cost = 8	Cost = 2	Cost = 8	Cost = 2	Subject ID[a]	Cost = 8	Cost = 2	Cost = 8	Cost = 2
1	1,020	921	538[b]	934	21	879	1,045	765	1,048
2	1,050	1,052	730	1,063	22	859	1,045	853	1,120
3	897	1,095	685	997	23	794	1,071	789	1,125
4	994	961	986	1,399[b]	24	829	1,052	840	1,093
5	632[b]	1,064	801	958	25	878	1,096	757	1,217
6	893	909	776	939	26	858	984	834	924
7	962	901	960	1,114	27	803	1,023	790	1,109
8	952	883	668	1,514[b]	28	819	1,030	791	1,135
9	952	1,024	849	965	29	842	1,025	792	1,168
10	916	993	992	1,059	30	815	1,022	892	1,070
11	872	942	799	879	31	873	1,005	777	1,142
12	968	998	880	880	32	782	994	799	1,075
13	968	958	660[b]	1,034	33	829	995	816	1,122
14	990	1,137	942	970	34	869	1,065	750	1,036
15	872	903	522[b]	1,103	35	871	962	832	1,054
16	925	994	848	916	36	795	1,061	789	1,141
17	926	988	893	1,119	37	815	983	839	1,006
18	1,029	988	951	1,129	38	794	1,048	823	1,094
19	930	1,040	854	1,029	39	820	1,010	789	1,068
20	1,069	1,093	936	1,132	40	846	1,136	787	1,218

[a]In each of the four treatments, subjects were assigned ID numbers from 1 to 20. Subjects with the same ID number in the same ρ scenario received the same demands. Each subject participated in only one treatment.

[b]Indicates cases where average orders do not exhibit pull-to-center.

the average order of each subject for all 40 rounds for all the four treatments. The data shows that the pull-to-center pattern in the orders are robust in that they are not driven by the aggregation of different segments of subjects, for example, one that orders close to the level predicted by the standard inventory model and another segment that deviates wildly in other directions. Out of the 160 subjects, only 6 did not exhibit pull-to-center behavior. Hence, the pattern of orders predicted by the behavioral model is robust at the individual subject level. Moreover, we did not detect any major shifts in ordering behavior across decision rounds. When we compare orders in the first 20 rounds with the last 20 rounds, we did not find any significant differences in orders in all treatment conditions except for the high-profit and $\rho = 0.8$ scenario. Here the orders in the last 20 rounds increased by an average of 32 units and 50 units in the centralized and decentralized cases, respectively. However, these order levels are still significantly lower than those predicted by the standard inventory model.

RESULT 3. *Because of the pervasive pull-to-center pattern in orders, actual average profits are lower than predicted by the standard inventory model.*

The results of the one-sample t-tests reported in Table 2 indicate that the actual average profits are significantly lower than the expected profits predicted by the standard inventory model in six out of the eight treatment-ρ conditions. For the other two conditions (centralized and decentralized in the low-profit condition when $\rho = 0.8$), although the average profit levels are not significantly different from those predicted by the standard inventory model, they are directionally lower. These results are not surprising given that actual orders deviate significantly from the predictions of the standard model. Overall, we find empirical support for Proposition 3.

We can also examine if the risk-pooling benefits of centralization survives the pull-to-center pattern in orders. To do so, we compare the profit outcomes across the two inventory structures while controlling for the effect of demand on profits by matching profits of subjects with the same ID across treatments. We first consider the $\rho = 0$ scenario, where centralization is predicted to yield the greatest profit difference. In the low-profit condition, profits under centralization are indeed higher ($t = 10.98$, $p = 0.000$). Furthermore, the percentage difference in actual profits is greater than the predicted difference of 64% ($t = 2.26$, $p = 0.024$) by the standard inventory model. In the high-profit condition, we also find that profit levels under centralization are higher ($t = 7.37$, $p = 0.000$). In terms of percentage difference, the predicted profit difference of 8% cannot be rejected at the 5% level ($t = 1.98$, $p = 0.062$). Hence, when demands across

stores are uncorrelated, centralizing the inventory structure does indeed produce higher profits despite the pull-to-center biases in orders.

In contrast, we do not find empirical support for the superiority of centralization for the $\rho = 0.8$ scenario. In the low-profit condition, despite the profits being directionally higher under centralization, the profit difference is not statistically significant ($t = 1.37$, $p = 0.186$). We obtain the same pattern of results for the high-profit condition ($t = 1.66$, $p = 0.112$). Note that these results are obtained with 1,600 observations in each test and after controlling for the effect of demands on profits. Moreover, we confirm that under our experimental design, centralization would indeed yield higher profits across both profit conditions if subjects were to order at the quantities predicted by the standard inventory model even though the predicted difference in the expected profits is small.[10] Hence, in this case, the biases in order quantities reduces the benefit of risk pooling.

3.3. Estimating the Behavioral Parameters

The results reported above show that the experimental data indicate support for the predictions of the reference-dependence model. To test the propositions more rigorously, we proceed to structurally estimate the parameter values of the reference-dependence model using the entire experimental data set (i.e., for all $S = 160$ subjects and all $T = 40$ rounds). In the centralized treatments, we assume that order quantities made by each subject s in round t is normally distributed with mean Q_r^{c*} given by Equation (6) and variance due to $(\tau^c)^2$. Similarly, under decentralization, orders are assumed to be normally distributed with mean q_r^{d*} given by Equation (4) and variance $(\tau^d)^2$. Note that $\mu = 1,000$ and $\sigma = 400$ in both Q_r^{c*} and q_r^{d*}. The noise parameters $(\tau^c)^2$ and $(\tau^d)^2$ are specific to each treatment and for each ρ scenario. The behavioral parameters δ_o and δ_u are specified to be common across all the treatments because these psychological costs affect all decision makers across both inventory structures and profit conditions. We estimate the full behavioral model and two nested models that correspond to Schweitzer and Cachon's model and the standard inventory model using maximum likelihood.

The results of the estimation are displayed in Table 5. The figures in the parentheses are the t-statistics of the parameter estimates with the standard errors clustered

---

[10] To check this, we compare the simulated profit outcomes across the two inventory structures assuming that subjects order the quantities predicted by the standard inventory model. Our test results confirm that the difference in profits between centralization and decentralization is statistically significant in both the low-profit ($t = 12.4$, $p = 0.000$) and the high-profit ($t = 14.3$, $p = 0.000$) conditions.

Table 5   Estimation Results of the Full Behavioral Model and Nested Models

Estimated behavioral parameters	(1) Full behavioral model	(2) Symmetric costs for leftovers and stockouts $\delta_u = \delta_o$	(3) No reference dependence (standard inventory model) $\delta_u = \delta_o = 0$
δ_u	6.52 (12.15)*	5.85 (6.26)*	— —
δ_o	9.96 (12.10)*	—	—
$\tau^c_{lp,\,p=0}$	606.3 (7.95)*	593.4 (7.02)*	734.7 (14.26)*
$\tau^c_{hp,\,p=0}$	579.9 (15.20)*	647.1 (14.60)*	903.8 (17.33)*
$\tau^d_{lp,\,p=0}$	181.6 (11.25)*	187.1 (9.36)*	235.5 (13.60)*
$\tau^d_{hp,\,p=0}$	241.4 (8.35)*	255.9 (11.84)*	369.8 (19.12)*
$\tau^c_{lp,\,p=0.8}$	412.6 (8.94)*	436.7 (9.30)*	708.0 (17.11)*
$\tau^c_{hp,\,p=0.8}$	445.8 (11.59)*	598.8 (12.32)*	1,199.9 (35.08)*
$\tau^d_{lp,\,p=0.8}$	97.7 (9.96)*	110.9 (9.94)*	171.8 (17.01)*
$\tau^d_{hp,\,p=0.8}$	181.7 (12.21)*	180.1 (12.72)*	295.6 (21.33)*
−LL	45,391	45,890	48,496
Wald test and *p*-value (test against full behavioral model)		$\chi^2(1) = 72.9$ $p = 0.000$	$\chi^2(2) = 154.1$ $p = 0.000$

Note. The figures in parentheses are the *t*-statistics.
*Indicates significance at the 5% level.

at the subject level to account for within-subject correlation in ordering decisions. Column (1) of Table 5 shows the parameter estimates and model fit of the reference-dependence model that incorporates different psychological costs of leftovers and stockouts. The results demonstrate that $\delta_o > \delta_u > 0$: every unit that the decision maker could have sold but was not in stock carries a psychological cost of 6.52 points ($p = 0.000$), whereas every unit that is paid for but not sold costs 9.96 points ($p = 0.000$). These estimates also confirm

Table 6   In-Sample Order Predictions (per Store) of the Behavioral and Standard Models

	$\rho = 0$		$\rho = 0.8$	
Treatments	Behavioral model	Standard theory	Behavioral model	Standard theory
Centralized				
Low-profit (cost = 8)	907 (−34)	832 (−109)	819 (−15)	690 (−144)
High-profit (cost = 2)	1,029 (+37)	1,168 (+176)	1,053 (+21)	1,310 (+277)
Decentralized				
Low-profit (cost = 8)	804 (−9)	663 (−150)	804 (−1)	663 (−142)
High-profit (cost = 2)	1,057 (0)	1,337 (+280)	1,057 (−41)	1,337 (+239)

Note. The numbers in parentheses represent the deviation in units from the actual average order.

that the parametric requirements that yield Proposition 1 are satisfied: the ratio of δ_u/δ_o is 0.65, which is greater than $(p - c_{lp})/(c_{lp} + s) = 0.25$ in the low-profit condition and smaller than $(p - c_{hp})/(c_{hp} + s) = 4$ in the high-profit condition of our experiment. Note that the psychological cost of a leftover, which has to be ordered and paid upfront, is about 1.53 times that of a stockout. As predicted in Proposition 2, this asymmetry in the psychological costs explains the stronger pull-to-center effects in the high-profit condition.

The results of the nested models are shown in columns (2) and (3) of Table 5. Column (2) displays the results of the model that does not allow for different psychological costs of leftovers and stockouts (i.e., $\delta_o = \delta_u$), which is mathematically equivalent to Schweitzer and Cachon's (2000) model, while column (3) represents the standard inventory model ($\delta_o = \delta_u = 0$). The Wald tests show clearly that the behavioral model that allows for reference dependence with asymmetric psychological costs of leftovers and stockouts explains the data better than these two models.[11] Table 6 shows that the in-sample predictions of orders based on the estimated behavioral model track the actual average orders much

[11] Because the observations are not independent at the subject level, we used the Wald test instead of the traditional likelihood ratio test (Wooldridge 2001).

Table 7 **Comparison of Actual and Predicted Total Profits (Point Earnings) of the Behavioral and Standard Models**

Treatments	$\rho = 0$			$\rho = 0.8$		
	Actual average	Standard inventory model	Behavioral model predictions	Actual average	Standard inventory model	Behavioral model predictions
Centralized						
Low-profit (cost = 8)	4,633	5,760 (+1,127)	5,583 (+950)	3,606	3,870 (+264)	3,544 (−62)
High-profit (cost = 2)	28,082	29,760 (+1,678)	29,076 (+994)	26,578	27,870 (+1,292)	26,609 (+31)
Decentralized						
Low-profit (cost = 8)	2,601	3,521 (+920)	3,167 (+566)	3,250	3,521 (+271)	3,167 (−83)
High-profit (cost = 2)	25,265	27,521 (+2,256)	26,153 (+888)	26,190	27,521 (+1,331)	26,153 (−37)

Notes. In the experiments, we use different point-earnings-to-cash-conversion ratios across the low-profit and high-profit treatments so that subjects will receive the same expected monetary payoff across the treatments if they order the quantities prescribed by the standard inventory model. The numbers in parentheses represent the deviations from the actual average point earnings.

more closely than the standard inventory model. Hence, our paper demonstrates that generalizing the standard inventory model by incorporating decision biases in subjects' decision making can be valuable because it can yield a much more predictive model of actual ordering behavior.

Finally, we compare the profit predictions of the estimated reference-dependence model with those of the standard inventory model. Table 7 contrasts the actual average profits with those predicted by the two models. This table shows that the behavioral model tracks the actual profits better in terms of absolute profit deviations in all four treatments and across the two ρ scenarios. This superiority carries over to the individual subject level. Under centralization, the behavioral model predicts average profit better than the standard inventory model in 27 and 32 out of 40 subjects in the low-profit ($Z = 2.21$, $p = 0.027$) and high-profit ($Z = 3.79$, $p = 0.001$) conditions, respectively. This is also true under decentralization: the behavioral model dominates in 26 and 29 out of 40 subjects for the low-profit ($Z = 1.90$, $p = 0.058$) and high-profit ($Z = 2.85$, $p = 0.004$) conditions, respectively.

4. Validation Experiments

The experimental results corroborate the predictions of the behavioral model that orders across both the centralized and decentralized inventory structures will exhibit pull-to-center, with a stronger bias in the high-profit conditions. Although our model is rooted in the well-established concept of reference dependence and the values of the behavioral parameters are reasonable, we recognize that there may be other models that can generate similar pull-to-center predictions. Hence, it is important to validate our behavioral model by showing that the psychological processes that we posit do indeed drive ordering decisions. Furthermore, a good strategy to validate our behavioral model would be to apply our understanding of the proposed psychological biases to alleviate or even eliminate the pervasive pull-to-center effects.

We conduct two follow-up experiments to meet these twin objectives. The basic idea behind the two experiments is to attempt to manipulate the relative salience of δ_u and δ_o in order to reduce the pull-to-center effects. It is also important to note that in our experiments, we do not seek to eliminate reference dependence, that is, make δ_u and δ_o as close to zero as possible (if it were at all possible). Rather, we assume that these biases due to reference dependence exist and try to increase the relative effect of one bias over the other to influence ordering decisions in the intended direction.[12] We conduct both validation experiments in the simplest setting a centralized inventory structure with $\rho = 0$ and examine both the low-profit and high-profit conditions. The detailed instructions of the experiment can be found in Appendix B.

4.1. Experiment 2A

In the first experiment, we influence the relative strength of the reference-dependence parameters δ_u and δ_o in the following ways. For the low-profit condition, our goal was to reduce overordering (leading to a reduction of pull-to-center) by increasing δ_o/δ_u.

---

[12] Note that our validation strategy can only be operationalized in a model with separate psychological costs of leftovers and stockouts.

Table 8 Recording Losses from Leftovers in the Low-Profit Condition Alleviates Pull-to-Center Bias

	Pull-to-center bias exists? (Test: Order per store greater than 832)			Pull-to-center bias alleviated? (Test: Two-sample test with data from main experiment)		
	Difference in mean order	*t*-stat.	Pull-to-center exists?	Difference in mean order	*t*-stat.	Pull-to-center alleviated?
Overall data	54.2	6.3	Yes	−54.4	−9.1	Yes
Individual subjects				Test by matching subjects with the same ID in the main experiment		
1	101	7.8	Yes	−87	−4.2	Yes
2	−2	−0.2	No	−220	−13.7	Completely[a]
3	43	5.9	Yes	−21	−0.8	No
4	60	3.8	Yes	−102	−4.1	Yes
5	132	7.0	Yes	332	8.3	No
6	17	2.8	Yes	−44	−2.8	Yes
7	−5	−1.2	No	−135	−6.3	Completely
8	61	3.6	Yes	−58	−2.5	Yes
9	80	7.7	Yes	−39	−2.1	Yes
10	33	6.1	Yes	−50	−2.1	Yes
11	67	7.4	Yes	28	1.6	No
12	41	5.1	Yes	−95	−4.6	Yes
13	37	6.7	Yes	−98	−6.5	Yes
14	98	6.7	Yes	−60	−3.5	Yes
15	79	8.4	Yes	39	2.1	No
16	−20	−1.5	No	−113	−6.6	Completely
17	69	8.1	Yes	−26	−1.2	No
18	73	6.4	Yes	−124	−4.0	Yes
19	35	4.9	Yes	−63	−5.1	Yes
20	84	4.2	Yes	−153	−6.0	Yes

[a]The pull-to-center bias is completely eliminated.

To achieve this, we increase the relative salience of δ_o by requiring subjects to compute and write down the number of leftover units and the profit loss from having those leftover units (by multiplying the number of leftover units by $c_{lp} = 8$) at the end of every decision round on a sheet of paper. Subjects were asked to write down "0" if there were no leftovers. Notice that we did not ask subjects to compute and record the costs of not meeting demand as we did not want to increase the salience of δ_u. Except for this change, the experimental instructions and procedure was identical to that of the main experiment described above. In the high-profit condition, we had the opposite objective of reducing underordering. Our aim was to reduce δ_o / δ_u through increasing the salience of δ_u. Specifically, we asked subjects to compute the units that they could have sold and the corresponding loss in profits from not meeting the demand (by multiplying the number of units that could have been sold by the profit margin $p - c_{hp} = 8$) at the end of every round. If there was no unmet demand, subjects were told to write down "0."

The results for the low-profit and high-profit conditions are reported in Tables 8 and 9, respectively. In each table, we report both the aggregate findings and the results for each subject ID. First, from the left panel of Table 8, observe that despite asking subjects

to record the profit losses from leftovers, the pull-to-center effect still exists—at the aggregate level, orders per store remain 54 units higher than the level of 832 units predicted by the standard inventory model ($\delta_o = \delta_u = 0$). We also find that 17 out of 20 subjects exhibited pull-to-center behavior in their ordering decisions. Next, we examine the more critical question of whether the pull-to-center bias is alleviated in this experiment by comparing the orders in this experiment with those from the same treatment in the main experiment (i.e., centralized, low-profit, and $\rho = 0$) where subjects were not asked to record the losses from having leftovers. We perform the subject-level analysis by conducting two-sample t-tests for subjects with the same ID across the two experiments (we are able to match subjects because subjects with the same ID received the same demands). The right panel of Table 8 shows that in aggregate, average orders decreased by 54 units when subjects were asked to write down the losses from leftovers ($t = -9.1$, $p = 0.000$). At the individual level, the pull-to-center bias is reduced for 15 out of 20 subjects. We obtain similar results for the high-profit condition where subjects were asked to write down their losses from stockouts. Notice from the left panel of Table 9 that subjects continue to underorder by an average of 103 units relative to the optimal level. Also, all subjects exhibit the

Table 9 Recording Losses from Stockouts in the High-Profit Condition Alleviates Pull-to-Center Bias

	Pull-to-center bias exists? (Test: Order per store less than 1,168)			Pull-to-center bias alleviated? (Test: Two-sample test with data from main experiment)		
	Difference in mean order	*t*-stat.	Pull-to-center exists?	Difference in mean order	*t*-stat.	Pull-to-center alleviated?
Overall data	−102.8	−13.8	Yes	73.1	4.1	Yes
				Test by matching subjects with the same ID in the main experiment		
Individual subjects						
1	−120	−25.6	Yes	127	6.5	Yes
2	−109	−4.6	Yes	7	0.3	No
3	−35	−5.7	Yes	38	3.0	Yes
4	−72	−6.6	Yes	136	5.2	Yes
5	−119	−5.3	Yes	−15	−0.4	No
6	−157	−7.2	Yes	102	2.8	Yes
7	−110	−14.9	Yes	167	9.7	Yes
8	−115	−27.9	Yes	171	14.2	Yes
9	−134	−12.0	Yes	10	0.4	No
10	−130	−14.9	Yes	45	3.0	Yes
11	−81	−8.9	Yes	146	6.1	Yes
12	−154	−17.5	Yes	16	0.9	No
13	−50	−4.5	Yes	160	11.2	Yes
14	−80	−9.1	Yes	−49	−2.7	No
15	−103	−9.6	Yes	162	5.7	Yes
16	−139	−12.4	Yes	36	3.2	Yes
17	−97	−5.1	Yes	84	3.8	Yes
18	−71	−7.5	Yes	110	5.1	Yes
19	−72	−10.8	Yes	56	2.0	Yes
20	−120	−5.8	Yes	−45	−1.4	No

pull-to-center bias. This is not surprising because our previous results show that the bias is stronger in the high-profit condition. More importantly, we examine if asking subjects to focus on losses from stockouts increases orders relative to the case when they were not asked to. The right panel of Table 9 shows that when subjects were asked to record the losses from not meeting demand, orders indeed increase by an average of 73 units ($t = 4.1$, $p = 0.000$). At the subject level, we find also that the pull-to-center bias is alleviated in 14 out of 20 cases.

4.2. Experiment 2B

In this second experiment, we manipulated the relative salience of δ_u and δ_o using a different mechanism. For the low-profit condition, we assessed a 5¢ cash penalty if there are leftovers at the end of every round. Denote e as the corresponding point penalty in the experiment that is equivalent to a 5¢ cash penalty. The expected utility for a decision maker with a penalty of e when facing leftovers is thus given by

$$EU_e(Q_r)$$
$$= \int_{-\infty}^{Q_r} [(p-s)x - (c_{lp}-s)Q_r - \delta_o(Q_r-x) - e] \cdot f(x)\, dx$$
$$+ \int_{Q_r}^{\infty} [(p-c_{lp})Q_r - \delta_u(x-Q_r)] \cdot f(x)\, dx. \quad (13)$$

In the high-profit condition, we awarded a small 5¢ bonus for having no stockouts in each round (this can

be also thought of as a bonus for meeting a customer service guarantee). Denote b as the corresponding bonus points in the experiment that is equivalent to a 5¢ cash bonus. The expected utility is thus given by

$$EU_b(Q_r)$$
$$= \int_{-\infty}^{Q_r} [(p-s)x - (c_{hp}-s)Q_r - \delta_o(Q_r-x) + b] \cdot f(x)\, dx$$
$$+ \int_{Q_r}^{\infty} [(p-c_{hp})Q_r - \delta_u(x-Q_r)] \cdot f(x)\, dx. \quad (14)$$

Solving the FOC, we get

$$F(Q_r^*) = \begin{cases} \dfrac{p - c_{hp} + \delta_u + b \cdot f(Q_r^*)}{p + \delta_u + \delta_o} \\ \quad \text{for the high-profit condition,} \\ \dfrac{p - c_{lp} + \delta_u - e \cdot f(Q_r^*)}{p + \delta_u + \delta_o} \\ \quad \text{for the low-profit condition.} \end{cases} \quad (15)$$

As can be derived from Equation (15), the effect of introducing the small penalty of 5¢ in the standard model where $\delta_u = \delta = 0$ is to decrease the optimal per-store order from 832 to 818.[13] Under the behavioral

---

[13] In this treatment, the conversion rate between experimental points and cash payment in cents is given by 0.0087. A cash penalty of 5¢ is therefore equivalent to $e = 5/0.0087 = 574$ points in the experiment.

Table 10   Assessing a Small Penalty for Leftovers in the Low-Profit Condition Alleviates Pull-to-Center Bias

	Pull-to-center bias exists? (Test: Order per store greater than 817.75)[a]			Pull-to-center bias less than predicted? (Test: Order per store smaller than 902)[b]		
	Difference in mean order	t-stat.	Pull-to-center exists?	Difference in mean order	t-stat.	Pull-to-center alleviated?
Overall data one-sample t-test	35	3.8	Yes	−49	−5.4	Yes
Individual subjects one-sample t-test						
1	59	5.8	Yes	−25	−2.4	Yes
2	6	1.0	No	−79	−14.2	Completely[c]
3	40	11.5	Yes	−44	−12.7	Yes
4	70	10.5	Yes	−14	−2.2	Yes
5	−15	−1.1	No	−100	−6.9	Completely
6	52	15.0	Yes	−32	−9.2	Yes
7	27	2.5	Yes	−57	−5.1	Yes
8	−23	−1.8	No	−107	−8.5	Completely
9	0.1	0.01	No	−84	−6.4	Completely
10	−20	−1.4	No	−104	−7.3	Completely
11	−21	−2.7	No	−105	−13.5	Completely
12	96	6.7	Yes	12	0.8	No
13	111	8.2	Yes	27	2.0	No
14	60	5.0	Yes	−24	−2.0	Yes
15	17	1.5	No	−67	−5.7	Completely
16	24	2.6	Yes	−61	−6.6	Yes
17	40	24.8	Yes	−45	−27.8	Yes
18	56	5.5	Yes	−28	−2.7	Yes
19	13	0.4	No	−71	−2.2	Completely
20	102	8.6	Yes	18	1.5	No

[a]The number 817.75 is the optimal order per store when there are no behavioral biases taking into account the penalty for leftovers.
[b]The number 902 is the predicted optimal order per store taking into account the penalty for leftovers and assuming that $\delta_u = 6.52$ and $\delta_o = 9.96$, which are estimated from the data in the first experiment.
[c]The pull-to-center bias is completely eliminated.

model, if the values of the behavioral parameters were to remain unchanged at $\delta_u = 6.52$ and $\delta_o = 9.96$, then the predicted order in the presence of the penalty would be 902 units compared to 907 units without the penalty. The key hypothesis is that the penalty would not only decrease orders because subjects respond to this additional pecuniary incentive (as predicted in both the standard and behavioral models), but that the penalty would also serve to increase the relative salience of δ_o. If this indeed occurs, then we would expect orders in the experiment to be less than 902.

Similarly, Equation (15) suggests that the presence of the $b = 5¢$ bonus has the effect raising per-store orders from 1,168 to 1,203 in the standard inventory model where $\delta_u = \delta = 0$ and from 1,015 to 1,038 in the reference-dependence model assuming that the values of the behavioral parameters remain unchanged. Again, the idea is that the bonus would also have the additional effect of increasing the relative salience of δ_u, so that orders would be greater than 1,038. Across both profit conditions, note that the pull-to-center bias would be attenuated if the respective

intended effects on δ_o and δ_u due to the introduction of the penalty/bonus are operative.

The results of this experiment are reported in Tables 10 and 11. The left panel of Table 10 shows that the pull-to-center bias persists in the orders of 12 out of 20 subjects in the low-profit condition when the small penalty for leftovers is levied. More critically however, orders are less than 902 in aggregate ($t = −5.4$, $p = 0.000$) and in 9 out of these 12 subjects. Interestingly, the pull-to-center bias is eliminated for 8 subjects with the penalty. These results suggest that the penalty indeed has an effect of increasing the relative strength of δ_o, which reduces the tendency to overorder. When the bonus is awarded in the high-profit condition, the left panel of Table 11 shows that all 20 subjects continue to underorder. Again however, the pull-to-center bias is alleviated as orders are higher than 1,038 by an average of 52 units ($t = 5.9$, $p = 0.000$), suggesting that the bonus also increases the salience of δ_u. At the individual level, we find that the pull-to-center bias is reduced in 70% of the subjects.

Table 11 **Awarding a Small Bonus for Meeting Demand in the High-Profit Condition Alleviates Pull-to-Center Bias**

	Pull-to-center bias exists? (Test: Order per store smaller than 1,202.5)[a]			Pull-to-center bias less than predicted? (Test: Order per store greater than 1,038.3)[b]		
	Difference in mean order	t-stat.	Pull-to-center exists?	Difference in mean order	t-stat.	Pull-to-center alleviated?
Overall data one-sample t-test	−112	−12.7	Yes	52	5.9	Yes
Individual subjects one-sample t-test						
1	−88	−7.0	Yes	77	6.1	Yes
2	−74	−10.3	Yes	90	12.5	Yes
3	−68	−12.4	Yes	96	17.6	Yes
4	−101	−8.1	Yes	63	5.1	Yes
5	−105	−41.4	Yes	60	23.7	Yes
6	−163	−5.9	Yes	1	0.1	No
7	−162	−10.6	Yes	2	0.1	No
8	−163	−15.0	Yes	2	0.2	No
9	−72	−4.1	Yes	92	5.3	Yes
10	−103	−12.1	Yes	61	7.2	Yes
11	−157	−17.3	Yes	7	0.8	No
12	−70	−4.8	Yes	94	6.5	Yes
13	−82	−5.4	Yes	82	5.4	Yes
14	−72	−3.7	Yes	92	4.7	Yes
15	−181	−17.5	Yes	−16	−1.6	No
16	−160	−14.9	Yes	5	0.4	No
17	−129	−20.4	Yes	35	5.5	Yes
18	−100	−14.5	Yes	64	9.3	Yes
19	−65	−4.7	Yes	99	7.1	Yes
20	−133	−10.2	Yes	32	2.4	Yes

[a]The number 1,202.5 is the optimal order per store when there are no behavioral biases taking into account the bonus for meeting demand.

[b]The number of 1,038.3 is the predicted optimal order per store taking into account the bonus and assuming that $\delta_u = 6.52$ and $\delta_o = 9.96$, which are estimated from the data in the first experiment.

5. Discussion and Conclusion

In this paper, we present a behavioral model that incorporates reference dependence to capture decision making in a multilocation newsvendor setting. We show experimentally that the model can explain the pervasive pull-to-center biases in orders across both centralized and decentralized inventory structures and the high- and low-profit conditions. Moreover, our model provides an intuitive psychological explanation for why the pull-to-center effect is stronger in the high-profit conditions—this is because decision-makers' psychological disutility for leftovers is stronger than that for stockouts. Our estimates indicate that the psychological aversion for having paid for something that turns out to have no value is 53% more painful than missing out on purchasing something that turns out to have value. Not only does our behavioral model track orders remarkably well, we also show that it beats out the standard inventory model in predicting profit outcomes. Finally, we conduct two validation experiments to show that the psychological posits of the behavioral model are operational. We do so by influencing the relative salience of the psychological cost of a leftover versus the psychological cost of a stockout to alleviate the pull-to-center bias.

The discussion in §2 has shown that our model is a mathematical generalization of Schweitzer and Cachon (2000) with $\delta_u = \delta_o = \delta$. We note from Table 5 that allowing for $\delta_o > \delta_u$ explains the experimental data better. To the best of our knowledge, this paper is the first to provide benchmark estimates of the relative weights of the disutility from incurring actual versus forgone losses. More importantly, the experiments in §4 provide support for the psychological drivers underlying our model—these experiments rely on the key premise that there are two different reference-dependence parameters and that we can increase the relative salience of one type of bias over the other.

Our paper also complements two other approaches that have been applied to study decision biases in newsvendor models. Su (2008) developed a quantal response framework that assumes that decision makers can make mistakes but that more costly mistakes occur less frequently. Our paper extends his approach by providing a psychological basis for why decision makers might appear to make errors. Moreover, we note that the quantal response model predicts that the modal order quantity will still be at that quantity prescribed by the standard inventory model. We do not

find this pattern in our experimental data.[14] Next, Bostian et al. (2008) showed that an adaptive learning model that incorporates memory, reinforcement, and probabilistic choice (Camerer and Ho 1999) tracks order patterns across decision rounds well. We did not find any significant pattern of learning in our experiments, which has 40 decision rounds. Our findings are similar to those of Bolton and Katok (2008) who allowed subjects to make orders for 100 decision rounds in a single-store newsvendor experimental setting. They reported that orders exhibited a gradual trend toward the optimal order (if at all) only after 30 rounds. We believe that although examining how decision makers learn over time is useful, it is equally, if not more important, to explain behavior in situations where the number of decision rounds is limited. This is because in many corporate settings, managers may not have the opportunity to "learn" through making multiple ordering decisions over time. Hence, a model that can explain ordering decisions in the "early stages" well, especially those decisions in more complex multilocation newsvendor settings, would be a useful contribution to the literature.

Our research is part of the emerging field of behavioral operations management where the central interest is to both understand and model actual human decision making in operations management areas. Our behavioral model is based on the well-established concept of reference dependence and nests the standard inventory model as a special case. The structural approach we have taken also allows us to formally quantify the extent of the psychological biases, which is one of the novel contributions of this paper.

There are many potential future research opportunities. One future research direction is to investigate the case where unmet demands can be back-ordered but there is a waiting cost experienced by customers. A natural question here is whether managers underestimate the waiting costs experienced by customers. Another relevant research direction is to allow product returns (Pasternack 1985). We suspect that allowing product returns may reduce δ_o because returning unsold units reduces the actual overage cost. It will also be interesting to extend the current model to allow endogenous retail prices (Petruzzi and Dada 1999, He et al. 2008). We suspect that such a change may alleviate the psychological disutility to leftovers and stockouts to a certain extent because the firm

can now adjust prices based on realized demands. However, we do not think that the biases will be completely eliminated because it may not be always optimal to set the retail price to equate orders with demand, so that there will still be leftovers and stockouts. Another future research direction is to investigate the effects of decision biases in a broader supply chain context, which can be very useful in generating insights on how supply chain partners can work together to achieve high system efficiency; see Croson and Donohue (2006), Cui et al. (2007), Lim and Ho (2007), Ho and Zhang (2008), and Wu and Loch (2008) for some nascent studies along this line of research. Finally, it will be interesting to further examine the efficacy of centralized inventory structures over decentralized ones for broader sets of demand relationships across different locations.

Appendix A

PROOF OF PROPOSITION 1. When the critical ratio in the standard model is larger than 0.5, i.e., when $(p - c_{hp})/(p - s) \geq 1/2$ or $p + s - 2c_{hp} \geq 0$, there is pull-to-center if the critical ratio in the model with reference dependence is smaller than that in the standard model. This leads to

$$\frac{p - c_{hp} + \delta_u}{p - s + \delta_u + \delta_o} < \frac{p - c_{hp}}{p - s}. \qquad (A1)$$

It is easy to show that Equation (A1) is equivalent to $\delta_u/\delta_o < (p - c_{hp})/(c_{hp} + s)$. The proof of Equation (8) follows similarly.   Q.E.D.

PROOF OF PROPOSITION 2. Let us prove the sufficient condition here, i.e., $\delta_o > \delta_u > 0$ $(\delta_u > \delta_o > 0)$ leads to stronger (weaker) pull-to-center effect in the high-profit condition compared to a symmetric low-profit condition. The proof of the necessary condition follows in a similar fashion. Suppose the critical ratio in the high-profit condition in the model without reference dependence is given by $(p - c_{hp})/(p - s) \geq 1/2$, and the critical ratio in the low-profit condition is given by $(p - c_{lp})/(p - s) < 1/2$. The symmetry between the two profit conditions leads to $(p - c_{hp})/(p - s) = 1 - (p - c_{lp})/(p - s)$, i.e.,

$$c_{hp} = p - c_{lp} + s. \qquad (A2)$$

The critical ratios in the model with reference dependence are given by $(p + \delta_u - c_{hp})/(p + \delta_u - s + \delta_o)$ in the high-profit

condition and $(p + \delta_u - c_{lp})/(p + \delta_u - s + \delta_o)$ in the low-profit condition. This leads to

$$\frac{p + \delta_u - c_{hp}}{p + \delta_u - s + \delta_o} = \frac{p + \delta_u - (p - c_{lp} + s)}{p + \delta_u - s + \delta_o} < 1 - \frac{p + \delta_u - c_{lp}}{p + \delta_u - s + \delta_o} \quad \text{(A3)}$$

for any $\delta_o > \delta_u > 0$. This leads to

$$\begin{cases} Q^{c*}_{r,hp} - n \cdot \mu < n \cdot \mu - Q^{c*}_{r,lp} & \text{in centralization,} \\ q^{d*}_{r,hp} - \mu < \mu - q^{d*}_{r,lp} & \text{in decentralization.} \end{cases} \quad \text{(A4)}$$

The symmetry between the two profit conditions also implies that

$$\begin{cases} Q^{c*}_{hp} - n \cdot \mu = n \cdot \mu - Q^{c*}_{lp} & \text{in centralization,} \\ q^{d*}_{hp} - \mu = \mu - q^{d*}_{lp} & \text{in decentralization.} \end{cases} \quad \text{(A5)}$$

Equations (A4) and (A5) lead to $\Omega^c_{hp} > \Omega^c_{lp}$ and $\Omega^d_{hp} > \Omega^d_{lp}$.

In a similar fashion, it can be proven that $\Omega^c_{hp} < \Omega^c_{lp}$ and $\Omega^d_{hp} < \Omega^d_{lp}$ if and only if $\delta_u > \delta_o > 0$.   Q.E.D.

PROOF OF PROPOSITION 3. The expected profit $E\Pi(Q)$ as a function of the ordering quantity Q ($Q = Q^d/n$ under decentralization or Q^c under centralization) for a retailer facing with a demand distribution with p.d.f. $h(x)$ is given by

$$E\Pi(Q) = \int_{-\infty}^{Q} [px + s(Q - x)] \cdot h(x)\, dx$$
$$+ \int_{Q}^{\infty} pQ \cdot h(x)\, dx - cQ, \quad \text{(A6)}$$

where $H(\cdot)$ is the focal distribution of demand that the retailer uses to make orders—under centralization, $H(\cdot) = G(\cdot)$ and under decentralization $H(\cdot) = F(\cdot)$. Standard steps lead to

$$E\Pi(Q) = (p - c)\mu_H - (c - s)(Q - \mu_H)$$
$$- (p - s)\sigma_H R\left(\frac{Q - \mu_H}{\sigma_H}\right), \quad \text{(A7)}$$

where μ_H and σ_H^2 are the mean and variance of the focal demand with distribution $H(\cdot)$ and $R(\tau) = \int_{\tau}^{\infty} (y - \tau) \cdot (1/\sqrt{2\pi})e^{-y^2/2}\, dy$. It is thus easy to show that the profit difference for two orders Q^* and Q_r^* is given by

$$E\Pi(Q^*) - E\Pi(Q_r^*)$$

$$= (p - c)\mu_H - (c - s)(Q^* - \mu_H) - (p - s)\sigma_H R\left(\frac{Q^* - \mu_H}{\sigma_H}\right)$$

$$- \left[(p - c)\mu_H - (c - s)(Q_r^* - \mu_H) - (p - s)\sigma_H R\left(\frac{Q_r^* - \mu_H}{\sigma_H}\right)\right]$$

$$= (c - s)(Q_r^* - Q^*) + (p - s)\sigma_H$$

$$\cdot \left[R\left(\frac{Q_r^* - \mu_H}{\sigma_H}\right) - R\left(\frac{Q^* - \mu_H}{\sigma_H}\right)\right]. \quad \text{(A8)}$$

Because the optimal order in the standard model Q^* is solved through FOC and such an optimal order is unique, any other order $Q_r^* \neq Q^*$ will lead to a lower profit, $E\Pi(Q_r^*) < E\Pi(Q^*)$.   Q.E.D.

Appendix B. Instructions

Instructions for Experiment 1
This is an experiment in decision making. If you follow the instructions carefully and make good decisions, you

could earn a considerable amount of money that will be paid to you in cash before you leave today. What you earn today partly depends on your own decisions and partly on chance. It is important that you do not look at the decisions of others, and that you do not talk, laugh, or make noises during the experiment. You will be warned if you violate this rule the first time. If you violate this rule twice, you will be asked to leave the room immediately and your cash earnings will be $0. There will be a total of 40 decision rounds in this experiment. In each round, you will make point earnings. The more points you earn, the more cash earnings you make. The decision steps and how you earn points in every round are described as follows:

In each round, you will act as a manager in charge of four retail stores: Store A, Store B, Store C, and Store D. All four retail stores make profits by selling products to consumers, but to be able to sell the products, you must first order the products from a supplier before the selling season begins. Your decision is to decide how many units of the product to order (call this decision "Order").

1. In each round, you will make a single order decision that represents the total order for all the four stores.

2. Each unit of the product that is sold to consumers retails for a price of 10 points.

3. The cost of each unit of the product that you order from the supplier is 8 points.

4. For each of the four stores, the consumer demand, or the number of units that consumers want to buy (call them Demand_A, Demand_B, Demand_C, and Demand_D for Stores A, B, C, and D, respectively) is uncertain at the time of your ordering decision. The actual values of Demand_A, Demand_B, Demand_C, and Demand_D will be known to you only after you make the decision of how many units to order.

5. For each of the four stores, the consumer demand has a distribution that is approximately normal, with an average (i.e., mean) demand of 1,000 units and a standard deviation of 400 units. The median and mode of the demand distribution is also 1,000 units. The smallest possible demand value is 0 and the consumer demand is in integer numbers. As an illustration of the demand distribution, Table B.1 shows the chance (in percentages) that demand will fall within a certain range.

6. In each round, after you have submitted your order decision, the computer program will generate the

Table B.1   Distribution of Demands for Each Store

Range of demand for each store	Percentage chance that demand will fall in this range
0 to 200	2.3
201 to 400	4.4
401 to 600	9.2
601 to 800	15.0
801 to 1,000	19.1
1,001 to 1,200	19.1
1,201 to 1,400	15.0
1,401 to 1,600	9.2
1,601 to 1,800	4.4
1,801 and above	2.3

actual values of Demand_A, Demand_B, Demand_C, and Demand_D.

The values of Demand_A, Demand_B, Demand_C, and Demand_D are independently generated according to the distribution described above. That is, the demand for a store does not affect the demand value that will be drawn for another store. The computer will generate the demand values for each of the four stores in the same manner in every round. Moreover, for each store, the past values of demand are independent of the future values of demand.

7. If the *Total Consumer Demand* (i.e., Demand_A + Demand_B+Demand_C+Demand_D) is less than the number of units you order, the total number of units that the four stores will sell is the *Total Consumer Demand*. In this case there will be some units that you ordered for the stores but cannot sell. These unsold units of the product have zero value to you.

8. Conversely, if the *Total Consumer Demand* is greater than the number of units ordered, the number of units sold in total will be the number of units you ordered. In this case there will be some consumer demand that cannot be met.

9. Depending on the actual number of units you order (*Order*) and the *Total Consumer Demand*, your point earnings in each round is given as follows:

If *Total Consumer Demand < Order*: *Point Earnings = Retail Price*(Total Consumer Demand) − (Cost per unit ordered)*Order = 10*(Total Consumer Demand) − 8*Order*.

If *Total Consumer Demand > Order*: *Point Earnings = Retail Price*(Order) − (Cost per unit ordered)*Order = 10*(Order) − 8*Order*.

The computer program will calculate your point earnings in every round.

10. Your actual cash earnings for each round will be 0.0087*(*Your Point Earnings* − 1,152). That is, every 1,000 points you earn above the number 1,152 earns you 8.7¢. Note that your cash earnings will be negative if your point earnings for that round fall below 1,152. Your total cash earnings in this experiment will be $4 plus your cash earnings for the 40 decision rounds (we will deduct negative cash earnings from your total earnings as well). The computer program will calculate your cash earnings for each round and your total cash earnings for all the rounds you have completed. We will pay you immediately at the end of the experiment.

Instructions for Experiment 2

In Experiment 2, we manipulated the relative strength of the psychological costs of leftovers and stockouts for the treatment setting where the inventory structure is centralized and the demands across stores are independent. The instructions are identical to those of the main experiment, except that the additional points are inserted following point number 10.

Experiment 2A (Recording Losses from Leftovers or Stockouts).

1. *Low-profit condition*: In every round, we ask that you compute and record down the "*Loss from Leftover Units*" in the sheet provided. This loss is incurred whenever your *Order* is greater than the *Total Consumer Demand*. This loss in point earnings is the *Number of Unsold Units* (i.e., *Order − Total Consumer Demand*) multiplied by the cost per unit ordered (which is equal to 8 points). Please write "0" if there is no loss in that round.

2. *High-profit condition*: In every round, we ask that you compute and record down the "*Loss from Not Meeting Consumer Demand*" in the sheet provided. This loss is incurred whenever your *Order* is less than the *Total Consumer Demand*. This loss in point earnings is the *Demand That Was Not Met* (i.e., *Total Consumer Demand − Order*) multiplied by the per unit profit that could have been earned (which is the price of 10 points minus the cost per unit ordered of 2 points = 8 points). Please write "0" if there is no loss in that round.

Experiment 2B (Penalty for Leftovers/Bonus for No Stockouts).

1. *Low-profit condition*: In addition to the cash earnings as described above, there will be a penalty of 5¢ (i.e., 5¢ will be deducted from your total cash earnings) if there are unsold units for that round, that is, when your *Order* is greater than the *Total Consumer Demand* for that round. After every round, we ask that you record down whether your *Order* is greater than the *Total Consumer Demand* in the sheet provided. Please write down "1" if this occurs and "0" if it does not.

2. *High-profit condition*: In addition to the cash earnings as described above, you will receive a bonus of 5¢ whenever you meet the *Total Consumer Demand* for that round, that is, when your *Order* is greater than or equal to the *Total Consumer Demand* for that round. After every round, we ask that you record down whether your *Order* is greater than or equal to the *Total Consumer Demand* in the sheet provided. Please write down "1" if this occurs and "0" if it does not.

References

Bolton, G. E., E. Katok. 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.

Bostian, A. J., C. A. Holt, A. M. Smith. 2008. The newsvendor pull-to-center effect: Adaptive learning in a laboratory experiment. *Manufacturing Service Oper. Management* **10**(4) 590–608.

Camerer, C. F., T.-H. Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica* **67**(4) 827–874.

Camerer, C. F., R. Hogarth. 1999. The effects of financial incentives in experiments: A review and capital-labor-production framework. *J. Risk Uncertainty* **19**(1–3) 7–42.

Camerer, C. F., L. Babcock, G. Loewenstein, R. Thaler. 1997. Labor supply of New York city cab drivers: One day at a time. *Quart. J. Econom.* **112**(2) 407–442.

Croson, R., K. Donohue. 2006. Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Sci.* **52**(3) 323–336.

Cui, T. H., J. S. Raju, Z. J. Zhang. 2007. Fairness and channel coordination. *Management Sci.* **53**(8) 1303–1314.

Eppen, G. D. 1979. Effects on centralization on expected costs in a multi-location newsboy problem. *Management Sci.* **25**(5) 498–501.

Frederick, S., N. Novemsky, J. Wang, R. Dhar, S. Nowlis. 2009. Opportunity cost neglect. *J. Consumer Res.* **36**(4) 498–501.

Genosove, D., C. Mayer. 2001. Loss aversion and seller behavior: Evidence from the housing market. *Quart. J. Econom.* **116**(4) 1233–1260.

He, C., J. Marklund, T. Vossen. 2008. Vertical information sharing in a volatile market. *Marketing Sci.* **27**(3) 513–530.

Ho, T.-H., J. Zhang. 2008. Designing pricing contracts for boundedly rational customers: Does the framing of the fixed fee matter? *Management Sci.* **54**(4) 686–700.

Ho, T.-H., C. F. Camerer, J.-K. Chong. 2007. Self-tuning experience weighted attraction learning in games. *J. Econom. Theory* **133**(1) 177–198.

Ho, T.-H., N. Lim, C. F. Camerer. 2006a. Modeling the "psychology" of consumer and firm behavior with behavioral economics. *J. Marketing Res.* **43**(3) 307–331.

Ho, T.-H., N. Lim, C. F. Camerer. 2006b. How "psychological" should economic and marketing models be? *J. Marketing Res.* **43**(3) 341–344.

Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47**(2) 263–291.

Kőszegi, B., M. Rabin. 2006. A model of reference-dependent preferences. *Quart. J. Econom.* **121**(4) 1133–1165.

Lim, N. 2010. Social loss aversion and optimal contest design. *J. Marketing Res.* **47**(4) 777–787.

Lim, N., T.-H. Ho. 2007. Designing price contracts for boundedly rational customers: Does the number of blocks matter? *Marketing Sci.* **26**(3) 312–326.

Pasternack, B. A. 1985. Optimal pricing and return policies for perishable commodities. *Marketing Sci.* **4**(2) 166–176.

Petruzzi, N. C., M. Dada. 1999. Pricing and the newsvendor problem: A review with extensions. *Oper. Res.* **47**(2) 183–194.

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

Smith, V. 1982. Microeconomic systems as a laboratory science. *Amer. Econom. Rev.* **72**(5) 923–955.

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

Thaler, R. 1980. Toward a positive theory of consumer choice. *J. Econom. Behav. Organ.* **1**(1) 39–60.

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

Wooldridge, J. 2001. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

Wu, Y., C. Loch. 2008. Social preferences and supply chain performance. *Management Sci.* **54**(11) 1835–1849.