MANAGEMENT SCIENCE
Vol. 59, No. 2, February 2013, pp. 436–451
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)
http://dx.doi.org/10.1287/mnsc.1120.1638

The Effect of Payment Schemes on Inventory Decisions: The Role of Mental Accounting

Li Chen, A. Gürhan Kök
Duke University, Durham, North Carolina 27708 {li.chen@duke.edu, gurhan.kok@duke.edu}
Jordan D. Tong
University of Wisconsin–Madison, Madison, Wisconsin 53706, jtong@bus.wisc.edu

Does the payment scheme have an effect on inventory decisions in the newsvendor problem? Keeping the net profit structure constant, we examine three payment schemes that can be interpreted as the newsvendor's order being financed by the newsvendor herself (scheme O), by the supplier through delayed order payment (scheme S), and by the customer through advanced revenue (scheme C). In a laboratory study, we find that inventory quantities exhibit a consistent decreasing pattern in the order of schemes O, S, and C, with the order quantities of scheme S being close to the expected-profit-maximizing solution. These observations are inconsistent with the expected-profit-maximizing model, contradict what a regular or hyperbolic time-discounting model would predict, and cannot be explained by the loss aversion model. Instead, they are consistent with a model that underweights the order-time payments, which can be explained by the "prospective accounting" theory in the mental accounting literature. A second study shows that the results hold even if all physical payments are conducted at the same time, suggesting that the framing of the payment scheme is sufficient to induce the prospective accounting behavior. We further validate the robustness of our model under different profit conditions. Our findings contribute to the understanding of the psychological processes involved in newsvendor decisions and have implications for supply chain financing and contract design.

Key words: behavioral operations; newsvendor; mental accounting; decision under uncertainty
History: Received September 20, 2010; accepted August 20, 2012, by Christian Terwiesch, operations management. Published online in *Articles in Advance* November 5, 2012.

1. Introduction

In the newsvendor problem, a decision maker chooses an inventory order quantity to meet a random future demand, with the goal to maximize the expected profit from selling the product. This model framework is commonly used for managing products with a short selling season and limited replenishment opportunities, such as fashion apparel and high-tech products. Because of lack of historical data and unexpected fluctuations in cost parameters, many newsvendor-type decisions in practice are made by humans based on subjective judgments (e.g., the fashion buying problem studied by Fisher and Raman 1996). Furthermore, even if the demand distribution and cost parameters are fully specified, human subjects are often observed deviating significantly from the optimal solution in experiments (Schweitzer and Cachon 2000). In this paper, we study how seemingly innocuous differences in the payment scheme can lead to significantly different inventory decisions in newsvendor experiments.

We define a *payment scheme* in the newsvendor problem as the amounts and times of payment transactions associated with the order quantity, the realized sales, and the leftover units. There are many payment schemes used in practice, designed for reasons including risk sharing, cash constraints, and price discrimination. Clearly, a change in payment scheme may affect the order decision for perfectly rational reasons (if the net profit is altered). But will such a change also induce a certain behavioral effect? If so, to what extent? Finding answers to these questions can help inform the design of payment schemes in supply chain financing and contracting.

The literature on *mental accounting* (see Thaler 1999 for a review), most notably applied to consumer choice behavior, describes how individuals perceive and evaluate multiple financial transactions. It suggests that individuals may mentally aggregate and/or segregate transactions systematically based on factors such as time or an uncertain event before making evaluations. In the newsvendor setting, the random demand event sets a natural boundary for payment transactions, both in terms of time and an uncertain event. Thus, we posit that individuals mentally segregate payment transactions in the newsvendor problem into two time buckets before and after the demand realization, which we call "order-time payments" and "demand-time payments," respectively.

Because of this segregation, we reason that altering transactions before and after the demand realization in a payment scheme (while keeping the net profit constant) will lead to significantly different order decisions.

To examine this behavioral effect, we consider three payment schemes in our paper. The first payment scheme is similar to a standard wholesale price contract: the newsvendor pays for the order quantity at the order time and receives revenue after the demand realization. We call this payment scheme O ("own financing") because the order is financed by the newsvendor's own operating capital. In the second payment scheme, the newsvendor's order payment is delayed until after the demand realization. As a result, there is no order-time payments; all transactions occur after the demand realization. We call this payment scheme S ("supplier financing") because suppliers often offer this kind of cost-based loans to their customers (such as trade credits). However, the newsvendor can also obtain a similar cost-based loan from a third party (such as a bank). In the third payment scheme, the newsvendor receives advanced revenue for the order quantity at the order time, but must refund the advanced payment for the leftover units after the demand realization. Thus, the order-time payments constitute a net profit for the inventory quantity ordered. We call this payment scheme C ("customer financing") because large customers sometimes provide these kinds of revenue-based loans to their suppliers (e.g., O'Sullivan 2007). Nevertheless, the newsvendor can also obtain a similar revenue-based loan from a third party (such as a bank) using the inventory as collateral.

We present four descriptive models to predict newsvendor ordering behavior. The first model assumes that individuals correctly aggregate the order-time and demand-time payments and choose the order quantity that maximizes the expected profit. The second model assumes that individuals are loss averse with respect to the order-time and demand-time payments separately. The third model assumes that individuals discount the demand-time payments due to time-discounting preferences. Finally, the fourth model assumes that individuals underweight the order-time payments due to a mental accounting effect called "prospective accounting," in which individuals fully account for transactions looking forward in time, but largely discount transactions looking backward in time (Prelec and Loewenstein 1998).

We conduct three experimental studies with human decision makers to empirically test the predictions of the above models. In Study 1, we set the overage cost equal to the underage cost. Thus, the expected-profit-maximizing solution is to order the median demand under all three payment schemes, which allows us to neutralize the pull-to-center effect (i.e., the deviation from the optimal solution toward the center of the distribution; see Schweitzer and Cachon 2000). Our results show that order quantities exhibit a consistent decreasing pattern in the order of payment schemes O, S, and C, with the order quantities of scheme S being close to the expected-profit-maximizing solution. This ordering behavior is inconsistent with the loss aversion model. It is also in the opposite direction with the prediction of the time-discounting model (including hyperbolic discounting, Laibson 1997). Rather, the observed ordering behavior is consistent with the prospective accounting model. In Study 2, we further isolate the mental accounting effect from the physical timing of payments by having all physical payment transactions conducted after the demand realization. We find that the same ordering pattern in Study 1 is sustained. This result shows that the framing of the payment scheme is sufficient to induce the prospective accounting behavior. Finally, in Study 3, we demonstrate that the prospective accounting behavior is robust under high- and low-profit conditions, though the magnitude of the order differences across payment schemes is influenced by the profit condition.

Perhaps the most surprising finding from our study is that the behavioral effect of payment scheme works in the opposite direction against time discounting. When the interest rate is significant and/or the capital constraint is binding, the behavioral effect may be dominated by the time-discounting effect due to the tangible financial costs. However, when capital constraint is not an issue and the interest rate is negligible (as in our experiments), loans based on purchase cost (scheme S) or projected revenue (scheme C) might inadvertently lower the retailer's order quantity relative to the case without the loan (scheme O). Therefore, to avoid any unintended consequences in practice, one should carefully evaluate the relative magnitude of these opposing effects when designing a payment scheme for supply chain financing and contracting.

Furthermore, our findings also provide a plausible explanation for the asymmetry of pull-to-center effect observed by Schweitzer and Cachon (2000) and Bolton and Katok (2008). These authors found that the pull-to-center effect is larger in the low-profit condition than in the high-profit condition. The framing of the payment scheme in their studies is similar to our scheme O. As a result, the asymmetry can be explained by the prospective accounting effect (see a discussion in Study 3). We further show that such asymmetry disappears under scheme S and is reversed under scheme C. Thus, the direction of the pull-to-center asymmetry is dependent on the framing of the payment scheme.

The rest of this paper is organized as follows. We provide a literature review in §2. We present the three stylized payment schemes and four decision models in §3. We present our experimental results in §4. We conclude with a discussion of the results and their managerial implications in §5.

2. Literature Review

There is a growing literature on behavioral operations management (see Bendoly et al. 2006 for a review). In this literature, researchers study how humans make operational decisions and how these decisions may differ from the rational decision. For example, Schweitzer and Cachon (2000) first found the pull-to-center effect in the newsvendor problem. Various influencing factors in newsvendor decisions are also investigated, such as decision heuristics (Bostian et al. 2008), the role of learning and feedback (Bolton and Katok 2008, Lurie and Swaminathan 2009), demand estimation biases (Feiler et al. 2012), psychological costs (Ho et al. 2010), and bounded rationality (Su 2008, Kremer et al. 2010). In a serial supply chain setting, Sterman (1989) and Croson and Donohue (2005, 2006) found that human subjects do not sufficiently account for the pipeline inventory and subsequently overreact to their inventory levels, contributing to the bullwhip effect. Loch and Wu (2008) examined social preferences in supply chain contracts. Ho and Zhang (2008) and Katok and Wu (2009) further investigated the effectiveness of risk-sharing contracts, which are closely related to our paper because the behavioral effect of the payment scheme may also play a role in determining the effectiveness of risk-sharing contracts.

Mental accounting has long been used to help understand the psychology behind choice behavior (Kahneman and Tversky 1979, Tversky and Kahneman 1981, Thaler 1980). It provides an explanation for many phenomena in human behavior that seem irrational—most notably in consumer choice behavior (e.g., Thaler 1985, Heath and Soll 1996), and also in other functional areas such as finance (Shefrin and Statman 1985) and accounting (Burgstahler and Dichev 1997). A consumer's payment is mentally coupled with the consumption because the two are linked by the consumer good. Shafir and Thaler (2006) found that the typical wine connoisseur thinks of her initial purchase of a case of wine as an investment, and later thinks of the wine as free when she drinks it, and so goes through the entire process never experiencing the pain of payment. Similarly, Prelec and Loewenstein (1998) found that people prefer to prepay for a vacation because they think that a prepaid vacation is more pleasurable than one that must be paid for after returning. This is because the payment is less painful if there is a future vacation to anticipate, and the vacation is more enjoyable if the payment has already been made. Gourville and Soman (1998) called the gradual reduction in relevance of past payments "payment depreciation." More generally, Prelec and Loewenstein (1998) called the mental accounting rule that fully recognizes future payments but largely writes off past payments "prospective accounting." We contribute to the mental accounting literature by applying these concepts to the newsvendor problem.

Our paper is also related to the literature on the interface of operations and finance. The payment schemes we consider are actually stylized versions of real practices. Payment scheme O corresponds to the standard wholesale price contract. Payment scheme S is similar in terms of the timing of the payments to trade-credit arrangements observed in practice. For example, retailers often delay their payments to their supplier, taking advantage of trade credits offered by the supplier (Peterson and Rajan 1997, Ng et al. 1999). The trade-credit terms can certainly affect a firm's optimal ordering policy (Haley and Higgins 1973, Gupta and Wang 2009, Song and Tong 2012). However, our scheme S is stylized because it does not reflect the lower interest rate benefits that such arrangements typically offer relative to bank loans. Payment scheme C is similar to receiving revenue in advance from the customer or financing inventory from an external party using inventory as a collateral. Small suppliers sometimes seek revenue-based loans by using their inventory as collateral. This has been gaining popularity in practice because of an increase in buyer-based supply chain financing solutions (see O'Sullivan 2007). This scheme is also stylized, as the amount financed (or the revenue advanced) in practice may only be part of the total selling value of the inventory investment. Besides the practices that are closely related to our three payment schemes, we note that there are other financial considerations that can also affect inventory decisions, such as asset-based financing (Buzacott and Zhang 2004) and capital constraints (Xu and Birge 2004, Babich and Sobel 2004, Xu and Zhang 2010).

3. Models of Newsvendor Decision Making

In the newsvendor problem, a decision maker chooses an order quantity q of a product to meet a future random demand D. Let $F(\cdot)$ denote the cumulative distribution function for the random demand. We assume that backlogs are not allowed (i.e., unmet customer demand is lost) and leftover inventory cannot be carried over to the subsequent period and has zero salvage value. The unit cost of the product is c, and the selling price is p (with $p > c$).

Table 1 Net Payments and Transaction Timing Under Different Payment Schemes

	Order-time payments	Demand-time payments	
Payment scheme	Per unit ordered	Per unit sold	Per unit leftover
O	$-c$	$+p$	0
S	0	$+(p-c)$	$-c$
C	$+(p-c)$	0	$-p$

Consider the following three payment schemes. (1) In payment scheme O (own financing), the newsvendor pays the cost c per unit at order time and receives a revenue p per unit sold after the demand realization. (2) In payment scheme S (supplier financing), the newsvendor pays nothing at order time; after the demand realization, she receives $p-c$ per unit sold and pays the external financing party (such as the supplier) c per unit leftover. (3) In payment scheme C (customer financing), the newsvendor receives $p-c$ per unit ordered at order time, but must refund p per unit leftover back to the external financing party (such as the customer) after the demand realization. The payment schemes are summarized in Table 1.

Next, we describe four models that differ in how the decision maker takes the payment scheme into account. We use the first normative model, which predicts the same order decision across payment schemes, as a benchmark. The other three models, inspired by behavioral decision-making and consumer behavior literature, predict different ordering behaviors across payment schemes. In what follows, we use the term "reward function" to denote how the individual evaluates payment transactions under a given decision model.

3.1. Expected-Profit-Maximizing Model

Let $R^i(q,D)$ denote the reward function given the quantity q and demand realization D under the payment scheme $i \in \{O,S,C\}$. If the decision maker correctly aggregates the order-time and demand-time payments and chooses the optimal quantity to maximize the expected profit, then $R^i(q,D)$ is simply the net profit given by

$$R^i(q,D)=\begin{cases}-cq+p\min(q,D) & \text{if } i=\text{O},\\ (p-c)\min(q,D)-c\max(q-D,0) & \text{if } i=\text{S},\\ (p-c)q-p\max(q-D,0) & \text{if } i=\text{C}.\end{cases}$$

The optimal order quantity that maximizes the expected profit is $q^i=\arg\max_q \mathbb{E}_D[R^i(q,D)]$.

Proposition 1. *The expected-profit-maximizing quantities under the three payment schemes are $q^O=q^S=q^C=q^*$, where $q^*=F^{-1}((p-c)/p)$.*

It is easy to verify Proposition 1 by noting that $R^O(q,D)=R^S(q,D)=R^C(q,D)$ for any q and D. The term $(p-c)/p$ is known as the critical fractile.

3.2. Loss Aversion Model

Prospect theory, introduced by Kahneman and Tversky (1979), assumes that individuals are loss averse with respect to a reference wealth, such as the current net worth. In the newsvendor problem, it is reasonable to assume that individuals may update their reference wealth after the order-time payments. Thus, individuals may be loss averse with respect to both the order-time and demand-time payments. Let $\pi_1^i(q)$ denote the net payment at order time, and let $\pi_2^i(q)$ denote the net payment after the demand realization under payment scheme $i\in\{O,S,C\}$. As in Schweitzer and Cachon (2000), we capture loss aversion using the utility function $U^l(x)=\{x \text{ if } x\ge 0;\ \lambda x \text{ if } x<0\}$ for $\lambda>1$. Then the newsvendor's total reward is

$$R^i(q)=U^l(\pi_1^i(q))+U^l(\pi_2^i(q,D))$$
$$=\begin{cases}-\lambda cq+p\min(q,D) & \text{if } i=O,\\ 0+U^l[(p-c)\min(q,D) & \text{if } i=S,\\ \qquad -c\max(q-D,0)] & \\ (p-c)q-\lambda p\max(q-D,0) & \text{if } i=C.\end{cases}$$

The following proposition compares the optimal order quantities $q^i=\arg\max_q \mathbb{E}_D[R^i(q,D)]$ to the expected-profit-maximizing solution q^*.

Proposition 2. *If the decision maker is loss averse and updates her reference wealth after order-time payments, then the optimal quantities under the three payment schemes are all less than the expected-profit-maximizing solution, i.e., $q^O<q^*$, $q^S<q^*$, and $q^C<q^*$.*

Proof. The critical fractiles for payment schemes O and C are $(p-\lambda c)/p$ and $(p-c)/\lambda p$, respectively, which are both less than the expected-profit-maximizing critical fractile. Schweitzer and Cachon (2000) proved the result for payment scheme S, because they evaluated the loss aversion model assuming all payments are made at the same time. □

3.3. Time-Discounting Model

The decision maker may prefer to receive benefits earlier and delay costs until later, which is also known as the time-discounted utility model (see Frederick et al. 2002 for a discussion on time discounting). The discounting may be due to the real interest rate or behavioral preferences. Under this model, the decision maker discounts the demand-time payments because they occur later. Let δ $(0<\delta<1)$ denote this discount factor. The reward function $R^i(q, D)$ can then be expressed as follows:

$$R^i(q, D) = \begin{cases} -cq + \delta p \min(q, D) & \text{if } i = \text{O}, \\ \delta(p-c)\min(q, D) - \delta c \max(q-D, 0) & \text{if } i = \text{S}, \\ (p-c)q - \delta p \max(q-D, 0) & \text{if } i = \text{C}. \end{cases}$$

The following proposition compares the optimal order quantities $q^i = \arg\max_q \mathbb{E}_D[R^i(q, D)]$ to the expected-profit-maximizing solution q^*.

Proposition 3. *Under the time-discounting model, the optimal quantities under the three payment schemes have the following relationship*: $q^O < q^S = q^* < q^C$.

Proof. The critical fractiles for payment schemes O, S, and C are $(\delta p - c)/\delta p$, $(p-c)/p$, and $(p-c)/\delta p$, respectively. Because $(\delta p - c)/\delta p < (p-c)/p < (p-c)/\delta p$, we have $q^O < q^S = q^* < q^C$. □

Note that because there are only two time points in the model setup, there is no difference between standard time discounting and hyperbolic time discounting (Laibson 1997).

3.4. Prospective Accounting Model

Thaler (1985) suggested that for a consumer, the payment and consumption in a transaction are not seen as a separate loss and a gain, respectively. Rather, the payment is mentally "coupled" with the thought of the associated consumption, and the consumption is "coupled" with the thought of the associated payment (Prelec and Loewenstein 1998). However, the strength of these two couplings are not equal, and are strongly dependent on the sequence of events. Specifically, individuals use a mental accounting rule called "prospective accounting," in which coupling is stronger when looking forward in time, but weaker when looking backward in time. The resulting phenomenon is consistent with underweighting whichever occurs first: the payment or the consumption. For example, consider how the prospective accounting rule applies to the case when payment precedes consumption, such as in a prepaid vacation. From the vantage point of the payment, the pain of payment is buffered because it is strongly coupled with the anticipated pleasure of the future vacation. From the vantage point of the vacation, the pleasure of the vacation is decoupled from the pain of the payment because it occurred in the past. Thus, in this case, the result of the prospective accounting is an overall underweighting of the pain of payment.

Instead of payment and consumption, the newsvendor simply has outgoing payments and incoming payments, which occur either before or after the demand realization. The random demand event sets a natural separation point for payment transactions before and after it because it involves the resolution of an uncertain event. Thus, although there are three possible payment transactions (per unit ordered, per unit sold, and per unit leftover), we posit that individuals mentally segregate these transactions into two time buckets before and after the random demand event as shown in Table 1. Similar to the consumer, we propose that, for the newsvendor, the order-time payments are coupled with the demand-time payments because they are connected by the number of units ordered. Thus, assuming outgoing payments are analogous to consumer payments (both are negative utilities) and incoming payments are analogous to consumption (both are positive utilities), we can implement the predictions of prospective accounting for our three payment schemes. Under payment scheme O, we assign an underweighting factor β $(0 < \beta < 1)$ to the order-time payments (which is the order cost). Under payment scheme C, we assign the underweighting factor α $(0 < \alpha < 1)$ to the order-time payments (which is the net profit from the quantity ordered). Under payment scheme S, no transactions occur at the order time, so there is no underweighting. The reward function $R^i(q, D)$ under the prospective accounting model is as follows:

$$R^i(q, D) = \begin{cases} -\beta cq + p\min(q, D) & \text{if } i = \text{O}, \\ (p-c)\min(q, D) & \text{if } i = \text{S}, \\ \quad -c\max(q-D, 0) & \\ \alpha(p-c)q - p\max(q-D, 0) & \text{if } i = \text{C}. \end{cases}$$

The following proposition compares the optimal order quantities $q^i = \arg\max_q \mathbb{E}_D[R^i(q, D)]$ to the expected-profit-maximizing solution q^*.

Proposition 4. *With prospective accounting, the optimal quantities under the three payment schemes have the following relationship*: $q^O > q^S = q^* > q^C$.

Proof. The critical fractiles for payment schemes O, S, and C are $(p - \beta c)/p$, $(p-c)/p$, and $\alpha(p-c)/p$, respectively. Because $(p-\beta c)/p > (p-c)/p > \alpha(p-c)/p$, we have $q^O > q^S = q^* > q^C$. □

We refer the reader to the appendix for a comparison of the above model with the consumer utility model of Prelec and Loewenstein (1998).

3.5. Summary

In this section, we have derived the predictions of four behavioral models that predict various ordering patterns under payment schemes O, S, and C. We provide a summary of the model predictions in Table 2. We note that the effects in each model need not be mutually exclusive. For example, in reality, time discounting and prospective accounting may exist simultaneously. Because they bias orders

Table 2 Summary of Model Predictions for Orders Under Payment Schemes O, S, and C

Newsvendor decision model	Prediction
Expected-profit-maximizing model	$q^O = q^S = q^C = q^*$
Loss aversion model	$q^O < q^*, q^S < q^*, q^C < q^*$
Time-discounting model	$q^O < q^S = q^* < q^C$
Prospective accounting model	$q^O > q^S = q^* > q^C$

in opposite directions, the resulting order would depend on which effect dominates (see §5 for a further discussion).

4. Newsvendor Experiments

In this section, we present three newsvendor experiments to examine the behavioral effect of payment scheme on inventory decisions. To isolate the behavioral effect of the payment scheme, we eliminate factors such as capital constraints and interest rates in our experimental designs (see §5 for a discussion of the impact of these factors). In the first study, we test whether ordering behavior can be described by the models presented in the previous section. In the second study, we test whether the framing of the problem is sufficient to induce differences in order decisions even if all actual payments are made at the same time. In the third study, we test the robustness of the model predictions under high- and low-profit conditions.

4.1. Study 1: A Simple Payment Scheme Experiment

4.1.1. Experimental Design. In Study 1, we test the three payment schemes O, S, and C (see Table 1) under parameters $c = \$1$, $p = \$2$ in a repeated newsvendor setting.

In each round, subjects roll three fair six-sided dice, the sum of which determines the demand for that round. Thus, demand is independent, identically distributed, and symmetric with a minimum value of 3, maximum value of 18, and mean value of 10.5. We choose to generate random numbers using three dice instead of a computer to facilitate participant understanding of the payment schemes through active demand generation and counting. The distribution of the sum of three dice is also well approximated by a normal distribution.

Recall that all payment schemes are equivalent in the sense that they produce identical total net profits or losses for any given ordering decision and demand realization. Furthermore, the actual overage cost and underage costs are equal at \$1 each, and the expected-profit-maximizing solution under all payment schemes is to order either 10 or 11 units every period. The newsvendor pull-to-center effect suggests that participants are biased toward the mean of the demand distribution, or 10.5.

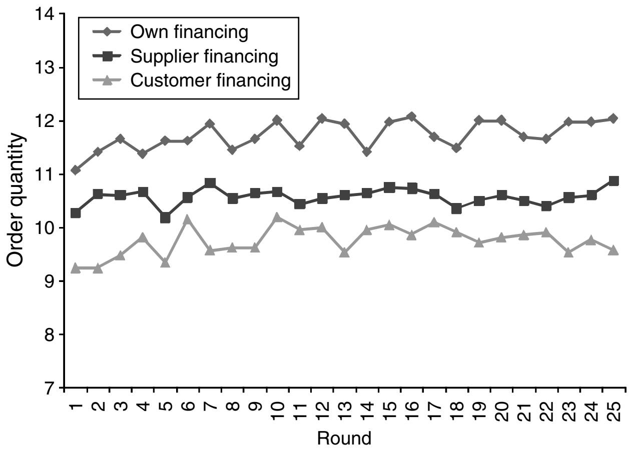
4.1.2. Methods. We recruited 99 undergraduate and graduate students from Duke University. Bolton et al. (2010) found that qualitatively students and managers perform similarly in the newsvendor problem. Croson and Donohue (2006) also found managers' and students' inventory decisions to be similar in a serial supply chain setting. Thus, we believe it is justifiable to use students as proxies for studying managerial behavior. The experimental conditions were assigned sequentially to the participants.[1] In exchange for their participation, participants received a minimum of \$5 plus a \$1 bonus for every 50 play dollars they had at the end of the game (each participant began with 100 play dollars). Participants earned anywhere from \$7–\$13 and took approximately 15 minutes to complete the experiment.

Participants were given an instruction sheet explaining the details of the game for the payment scheme to which they were assigned (see the supplemental material, available from the authors upon request). Instructions were also read out loud by a research assistant before beginning play. Participants were told that they would be selling "widgets" (represented by poker chips) and that customer demand for the widgets in a given round was represented by the sum of the rolling of three standard dice. Each participant interacted one on one with a research assistant, who facilitated payment transfers and recorded ordering decisions and dice rolls. A participant decided an order quantity vocally, placed that many poker chips into the "store" (represented by a square drawn on an index card), and made appropriate payment transfers. Then, the participant rolled the three dice, determined how many units were sold and/or leftover, and again made appropriate payment transfers. Finally, the participant removed all chips from the store to begin the next round.

Payment transfers were conducted in the form of play paper currency in denominations of 1, 5, and 10. All payments to the participant were conducted by the research assistant, whereas all payments from the participant were conducted by the participant. Appropriate payment transactions occurred immediately following the ordering decision and immediately following demand realization. The participant

[1] We conducted this study in two parts. In the first part, we randomly assigned one payment scheme, O or S, to each subject (57 subjects). In the second part, we randomly assigned one payment scheme, S or C, to each subject (42 subjects). Each subject completed the experiment under only one payment scheme. Although ideally we would have randomly assigned participants across all three payment schemes, we found no significant differences between the two repetitions of condition S, and therefore aggregated the data for analysis. This yielded 29 subjects for condition O, 49 for condition S, and 21 for condition C.

Figure 1 Average Order Quantities in Each Round in Study 1 Under Schemes O, S, and C



also moved the poker chips and rolled the dice themselves, which facilitated their understanding of the process. Game play was for 25 rounds, after which a follow-up written question was administered: "If you could play the game again choosing only one order quantity, what number would you choose?" Finally, two written comprehension questions were administered at this time: "What is the minimum demand possible you can roll with three dice?" and "What is the maximum demand possible you can roll with three dice?"

4.1.3. Results. All 99 participants completed the study. One participant in the S condition incorrectly answered both comprehension questions and also made multiple orders of more than 18, and was therefore removed from the analysis (though all results hold when included). The resulting average ordering decisions in each round are shown in Figure 1, and a summary of our main results can be found in Table 3.

We conducted a repeated measures generalized linear model to analyze the 25 inventory order decisions under each payment scheme. We found that payment scheme significantly affected ordering behavior ($F(2, 95) = 18.88, p < 0.0001$). Specifically, we found that orders were highest under payment scheme O and lowest under payment scheme C. To test these differences, we conducted planned contrast tests. These tests showed that all three differences were significant: orders under O were significantly greater than orders under S ($F(1, 95) = 18.10$, $p < 0.0001$), orders under S were significantly greater than orders under C ($F(1, 95) = 7.46, p = 0.0075$), and orders under O were significantly greater than orders under C ($F(1, 95) = 35.83, p < 0.0001$). As Table 3 shows, these same trends are present in the first ordering decision (which is not confounded by experience or feedback), the average order quantity, and the follow-up question. However, not all differences were significant. Specifically, the differences appear to be more significant for the average orders and the follow-up question than for the first ordering decision (see a discussion on this in the summary below).

Table 3 Mean and Standard Deviations of Order Quantities, and Significance Tests for Differences Between Payment Schemes in Study 1

Payment scheme	Mean order quantity (standard deviation in parentheses): Average over 25 rounds	Round 1	Follow-up question	N
O	11.728 (1.392)	11.069 (3.390)	11.759 (1.766)	29
S	10.573 (1.031)	10.271 (2.210)	10.448 (1.234)	48
C	9.749 (1.058)	9.238 (2.406)	9.571 (1.207)	21
		Contrast tests		
$q^O - q^S$	1.155***	0.798	1.311***	
$q^S - q^C$	0.824**	1.033	0.877*	
$q^O - q^C$	1.979***	1.831*	2.188***	

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

We also compared the average orders with the mean of the demand distribution, 10.5, because both the expected-profit-maximizing criterion and the pull-to-center effect predicted orders near mean demand. We found that average orders under O were significantly greater than mean demand ($t(28) = 4.751, p < 0.001$), and average orders under C were significantly less than mean demand ($t(20) = -3.256, p = 0.004$), whereas average orders under S were not significantly different from mean demand ($t(47) = 0.493$, $p = 0.624$).

The actual demands generated by rolling the three dice were relatively consistent with the theoretical predictions. The means were 10.739, 10.557, and 10.764, under O, S, and C, respectively. Also, all participants (except the one eliminated participant in condition S) correctly answered 3 and 18 for the minimum and maximum possible demand that could be generated by rolling the three dice.

There was no significant difference in the overall ordering levels over time (Wilks' lambda $= 0.701, F(24, 72) = 1.28, p = 0.212$). In other words, there was no main effect for round. We also found no significant interaction between payment scheme and experience gained as more rounds were played (Wilks' lambda $= 0.614, F(48, 144) = 0.83, p = 0.774$).

Summary. Study 1 establishes that payment schemes have a significant effect on ordering behavior in the newsvendor problem. We found that order decisions can be higher or lower than the expected-profit-maximizing decision depending on the payment scheme, which is inconsistent with the expected-profit-maximizing model. Specifically, we found that orders exhibit a consistent decreasing pattern in the order of schemes O, S, and C, with the order quantities of scheme S being close to the expected-profit-maximizing solution. These results are inconsistent with the loss aversion model and the time-discounting model. Rather, they are consistent with the prospective accounting model. Moreover, the differences appear to not only be robust over 25 rounds, but actually more significant over time. This suggests that the feedback individuals use to inform their future ordering quantity is also subject to the prospective accounting effect, making the order deviation robust over 25 rounds of experience (see §5 for a brief discussion of recency and anchoring and adjustment).

Structural Parameter Estimates. To provide further validation of the prospective accounting model, and to obtain estimates for its parameters, we estimated β and α using three structural estimation techniques. Note that the prospective accounting model reduces to the expected-profit-maximizing model when $\beta = \alpha = 1$. Therefore, we can validate the model fit of the prospective accounting model against the expected-profit-maximizing model by testing whether β and α are less than 1.

Our first two structural estimation approaches follow the N1 and N2 models of Olivares et al. (2008). The third approach we provide is a hybrid of N1 and N2, which we call NH. The differences between each approach lie in how they account for the variability in the observed order quantities. The N1 model attributes all of the order variability to a noisy underweighting factor, and uses an ordinary least squares regression to estimate the underlying underweighting factor. On the other hand, the N2 model attributes all of the order variability to errors in the order quantity (i.e., a "trembling hand") and uses a nonlinear least squares regression to estimate the underweighting factor. Finally, we also provide a hybrid approach that we believe is reasonable in our problem setting. The NH model attributes some of the variability to a noisy underweighting factor and some of the variability to a "trembling hand." It assumes that each participant has a unique underweighting factor (so that there is heterogeneity in underweighting factors across participants), but that the differences between order quantities across rounds for the same participant are due to a "trembling hand." We approximate the demand distribution with a normal distribution with mean 10.5 and standard deviation 2.958, matching the mean and standard deviation of the discrete demand distribution of the sum of three dice. The resulting estimates for β and α (denoted with $\hat{\beta}$ and $\hat{\alpha}$) are reported in Table 4.

The results from the three estimate models are relatively consistant, with the estimates from the hybrid model falling in between those of the N1 and N2 models. In other words, in this experiment we found that, on average, individuals' orders are consistent with taking into account only about 70% of payments that occur at the order time when costs precede revenues (payment scheme O), and only about 80% of payments that occur at the order time when revenues precede costs (payment scheme C). For example, under payment scheme O, this suggests that an individual who orders 10 units at \$1 each perceives the \$10 cost as if it were only about \$7.

Table 4 **Parameter Estimates for Study 1**

Parameter	Model N1	Model N2	Model NH
$\hat{\beta}$	0.7479 (0.0153)**	0.6780 (0.0192)**	0.7125 (0.0109)**
$\hat{\alpha}$	0.8231 (0.0154)**	0.7995 (0.0185)**	0.8137 (0.0110)**

Note. Standard errors are in parentheses.
**$p < 0.01$.

Expected Profits. We also calculated the expected profits given each ordering decision of each participant. Rather than using actual profits (which is an outcome-based measure), we use the expected profit measure because it captures the participants' decision efficiency. Expected profits were significantly affected by payment scheme ($F(2, 95) = 5.65$, $p = 0.005$). Average expected per-round profits (standard deviations in parentheses) by condition were 7.461 (0.696), 7.805 (0.240), and 7.746 (0.348) for O, S, and C, respectively. It is not surprising that the expected profits were highest under payment scheme S, because the average order quantity under S was closest to the expected-profit-maximizing quantity. Contrast tests show that all differences between conditions are significant differences at the $p < 0.05$ level except the difference between S and C.

One might suggest that these differences in profit are not extremely large (the expected per-round profits are 4.61% greater under S than under O). This is due to the fact that the expected profit function is relatively flat near the optimal solution. However, other operational metrics are not as flat around the optimal solution. For example, the supplier's revenue is the wholesale price times the newsvendor's order. Thus, the supplier's average per-round revenue is 20.29% greater under O than under C (on average, the supplier sells 11.728 units to the newsvendor versus 9.749). The differences between payment schemes also impact customer service. We calculated the customer's expected in-stock rate for each ordering quantity. This analysis shows that the average expected per-round in-stock rates are 0.675 (0.122), 0.564 (0.116), and 0.462 (0.120) for O, S, and C, respectively.

4.2. Study 2: A Payment Scheme Experiment with Same Payment Timing

4.2.1. Experimental Design. The purpose of Study 2 is to test whether we can achieve similar results to Study 1 by manipulating only the framing of the payment scheme (i.e., when and how payments are determined), while eliminating the difference in the actual timing of payments transactions. According to Prelec and Loewenstein (1998), by knowing the size of a payment before an uncertain event, a consumer can mentally impute that payment even if the actual time the payment transaction occurs later. Study 2 investigates whether a similar phenomenon exists for the newsvendor. It tests whether the framing of payment scheme in the newsvendor problem can induce individuals to mentally set aside payments at the order time, even if all payment transactions are actually conducted after the demand realization.

Study 2 implements nearly the same design as Study 1 with the following main exception: all payments are postponed to the end of each round (i.e., conducted after the demand realization), even if some payments are determined at the time of the ordering decision. We will refer to these three payment schemes as payment frames, and denote our conditions with an overbar: $\bar{\text{O}}, \bar{\text{S}}$, and $\bar{\text{C}}$. Because all payments are delayed until after the demand realization, there are no real differences between payment schemes in the actual financial position over time. However, payment frame $\bar{\text{O}}$ permits individuals to mentally set aside some cost at the order time, whereas payment frame $\bar{\text{C}}$ encourages individuals to mentally set aside some benefit at the order time. Thus, we expect results in Study 2 to be similar to those in Study 1.

4.2.2. Methods and Results. We recruited 57 undergraduate and graduate students from Duke University. The methods were the same as Study 1 except for the following differences. First, as described above, all payments were conducted at end of each round, after the demand realization. Second, the specific wording of the payment schemes varied slightly from Study 1: the financial transactions after demand realization were described in the sequence of per unit leftover and per unit sold, rather than per unit sold and per unit leftover as in Study 1 (see the supplemental material). This slight difference in wording, although unintentional, introduces an additional test for our two-period mental segregation model. An observation of the same behaviors in Studies 1 and 2 will imply the transaction sequence after the demand realization has little effect on individual's ordering decisions, which corroborates the assumption that individuals only segregate transactions before and after demand realization. Third, in addition to the comprehension questions asked in Study 1, at the end of the experiment we also asked each participant the question, "What do you think is the long-run average demand generated by rolling three dice?"

Table 5 Mean and Standard Deviations of Order Quantities, and Significance Tests for Differences Between Payment Frames in Study 2

Payment frame	Mean order quantity (standard deviation in parentheses): Average over 25 rounds	Round 1	Follow-up question	N
$\bar{\text{O}}$	11.648 (1.721)	10.800 (2.745)	11.900 (1.518)	20
$\bar{\text{S}}$	10.671 (1.020)	10.889 (2.055)	10.611 (0.850)	18
$\bar{\text{C}}$	9.665 (1.159)	9.474 (2.144)	9.632 (1.165)	19
Contrast tests				
$q^{\bar{O}} - q^{\bar{S}}$	0.977*	−0.089	1.289**	
$q^{\bar{S}} - q^{\bar{C}}$	1.006*	1.415	0.979*	
$q^{\bar{O}} - q^{\bar{C}}$	1.983***	1.326	2.268***	

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

All 57 participants completed the study and were included in the following analyses. A summary of our findings can be found in Table 5, and the average ordering decisions are shown in Figure 2. As in Study 1, the repeated measures generalized linear model showed that payment frame significantly affected ordering behavior ($F(2,54) = 10.94, p < 0.0001$). Orders under $\bar{\text{O}}$ were significantly greater than orders under $\bar{\text{S}}$ ($F(1,54) = 5.17, p = 0.0270$). Orders under $\bar{\text{S}}$ were significantly greater than orders under $\bar{C}$ ($F(1,54) = 5.34, p = 0.0246$), and orders under $\bar{\text{O}}$ were significantly greater than orders under $\bar{\text{C}}$ ($F(1,54) = 21.89, p < 0.0001$). Comparing average orders to the mean demand of 10.5, we again found that average orders under $\bar{\text{O}}$ were significantly greater than mean demand ($t(19) = 2.98, p = 0.007$), and average orders under $\bar{\text{C}}$ were significantly less than mean demand ($t(18) = -3.14, p = 0.006$), whereas average orders under $\bar{\text{S}}$ were not significantly different from mean demand ($t(17) = 0.80, p = 0.436$).

Even though formal comparisons between Studies 1 and 2 would not be appropriate because they are run at different times, we highlight the similarities

Figure 2 Average Order Quantities in Each Round in Study 2 Under Schemes O, S, and C

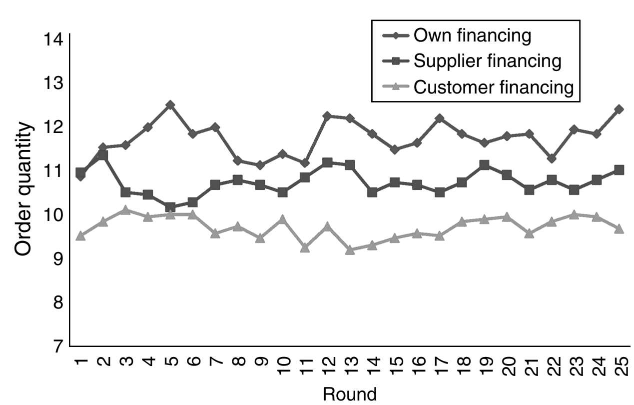


Table 6 Parameter Estimates for Study 2

Parameter	Model N1	Model N2	Model NH
$\hat{\beta}$	0.7555 (0.0205)**	0.6979 (0.0252)**	0.7389 (0.0861)**
$\hat{\alpha}$	0.8019 (0.0167)**	0.7778 (0.0187)**	0.7942 (0.0651)**

Note. Standard errors are in parentheses.
**$p < 0.01$.

and differences between the results in the two studies. Overall, Table 5 demonstrates ordering behavior remarkably similar to the results in Study 1, as given in Table 3. As in Study 1, orders did not significantly change over time (Wilks' lambda $= 0.472$, $F(24, 31) = 1.44$, $p = 0.1672$). We also found no significant interaction between payment frame and round (Wilks' lambda $= 0.309$, $F(48, 62) = 0.46$, $p = 0.4520$). On the other hand, in Study 2 there are no significant differences in the order quantities in round 1. Also, the overall significance of contrasts seem to be slightly less in Study 2 than in Study 1 for the average ordering quantities (two less contrasts are significant at the $p < 0.01$ level) and for the follow-up question (one less contrast is significant at the $p < 0.01$ level).

All participants correctly answered 3 and 18 for the minimum and maximum possible demand that could be generated by rolling three dice. Participant estimates of the long-run average demand with three dice revealed some variation, but were not significantly affected by condition ($F(2, 54) = 0.34$, $p = 0.714$). Their average estimates by condition were 10.35 (0.96), 10.25 (0.65), and 10.45 (0.47) for $\bar{\text{O}}$, $\bar{\text{S}}$, and $\bar{\text{C}}$, respectively (standard deviations in parentheses). The t-tests show that these estimates are also not significantly different from 10.5 (for $\bar{\text{O}}$, $t(19) = -0.7$, $p = 0.494$; for $\bar{\text{S}}$, $t(17) = -1.64$, $p = 0.1197$; for $\bar{\text{C}}$, $t(18) = -0.49$, $p = 0.630$). The actual mean demands were 10.787, 10.679, and 10.670, under $\bar{\text{O}}$, $\bar{\text{S}}$, and $\bar{\text{C}}$, respectively.

For Study 2, we followed the same procedure to estimate the β and α parameters as we did in Study 1. The results are reported in Table 6. The estimates are almost identical to what we observed in Study 1, which again confirms that payment framing is sufficient to induce underweighting consistent with the prospective accounting model.

The expected per-round profits in Study 2 also closely resembled the results from Study 1. Expected profits were significantly affected by payment frame ($F(2, 54) = 4.50$, $p = 0.016$). Average expected profits by condition were 7.402 (0.657), 7.824 (0.262), and 7.735 (0.341) for $\bar{\text{O}}$, $\bar{\text{S}}$, and $\bar{\text{C}}$, respectively. Contrast tests show that all differences between conditions are significant differences at the $p < 0.05$ level except between $\bar{\text{S}}$ and $\bar{\text{C}}$.

Summary. Study 2 establishes that the framing of the payment scheme can have a significant effect on order decisions even if the schemes have no differences in the actual timing of payments. This result provides further evidence for the mental accounting effect of payment schemes: by knowing the size of payments before the demand realization, individuals mentally set aside those payments and at the same time apply the prospective accounting rule to reach their order decision. In fact, we actually observed several participants who at the order time physically set aside or held in-hand the order-time payments, even though the actual transactions were not to be conducted until after the demand realization.

4.3. Study 3: Payment Scheme Experiments with High- and Low-Profit Conditions

4.3.1. Experimental Design and Hypotheses. In Study 3, we implement two repeated newsvendor experiments to test the effect of payment schemes for products with two different profit margins. The high-profit condition is conducted for a product with parameters $c = \$1$, $p = \$4$, which implies an actual overage cost of \$1 and an actual underage cost of \$3. Under the expected-profit-maximizing model, this yields a critical fractile of 75%. The low-profit condition is conducted for a product with parameters $c = \$3$, $p = \$4$, which implies an actual overage cost of \$3 and actual underage cost of \$1. Under the expected-profit-maximizing model, this yields a critical fractile of 25%. Within each high- and low-profit condition, we again test payment schemes O, S, and C. Because in practice payments are usually made when they are determined, we use the payment schemes in Study 1. One can substitute the appropriate values of c and p into Table 1 to obtain a description of the payment schemes for the high- and low-profit conditions.

For all payment schemes O, S, and C, the expected-profit-maximizing solution is 13 for the high-profit condition and 8 for the low-profit condition. The pull-to-center effect predicts that individuals are biased toward 10.5, the center of the distribution, causing actual orders to be somewhere between 13 and 10.5 for the high-profit condition, and somewhere between 8 and 10.5 for the low-profit condition. Nevertheless, the pull-to-center effect still predicts no difference between the payment schemes. Thus, although Study 3 does not allow us to determine the relative magnitude of deviations from the expected-profit-maximizing solution are whether due to the pull-to-center effect or due to payment schemes, it provides a test of robustness of the inequality predictions in Table 2 across high- and low-profit conditions.

4.3.2. Methods. We recruited 130 undergraduate and graduate students from Duke University—70 for the high-profit condition and 60 for the low-profit condition. The three payment schemes were assigned sequentially to the participants within each condition. In exchange for their participation, participants received a minimum of $5, with a bonus based on how much play money they earned in the game. In the high-profit condition, participants earned a $1 bonus for every 100 play dollars they had at the end of the game (each participant began with 100 play dollars). In the low-profit condition, participants earned a $1 bonus for every 50 play dollars they had at the end of the game (each participant began with 150 play dollars). For Study 3, each participant played the game for 20 rounds. In all respects except for the payment scheme parameter changes and the reduced number of rounds, the experimental design and methods were the same as in Study 1.

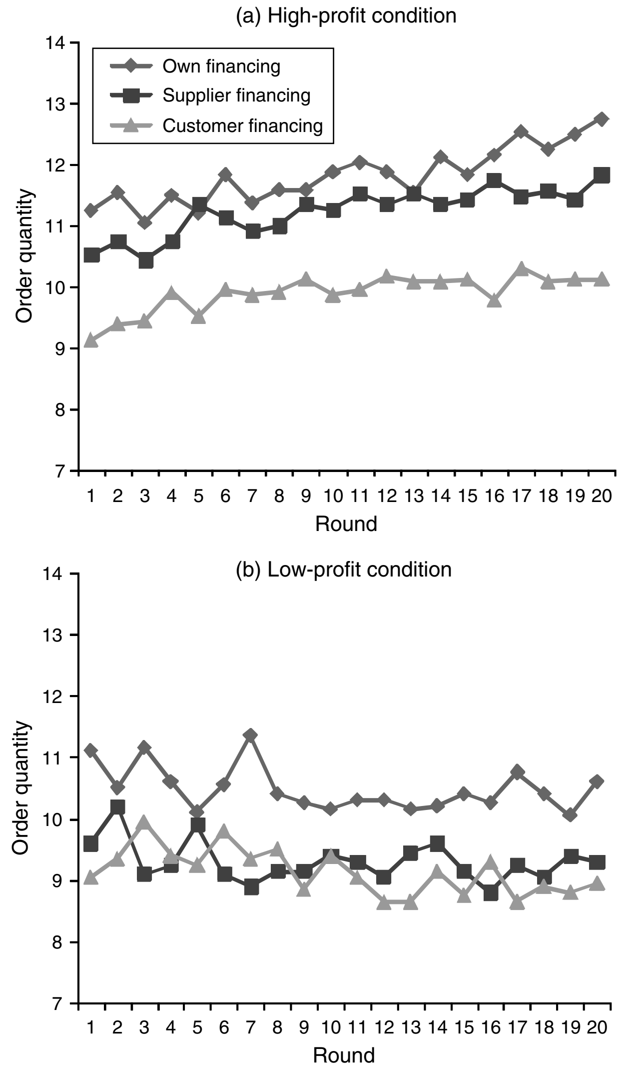
4.3.3. Results. All 130 participants completed the study and were included in the analyses. The resulting average ordering decisions for each round are shown in Figure 3, and a summary of our findings can be found in Table 7.

High-Profit Condition. For the high-profit condition, the repeated measures generalized linear model showed that payment scheme significantly affected ordering behavior ($F(2, 67) = 18.61$, $p < 0.0001$). We again found that average orders were highest under payment scheme O and lowest under payment scheme C. Follow-up planned contrasts showed that some of the differences between conditions were significant, whereas others were not. Orders under O were not significantly greater than orders under S ($F(1, 67) = 3.34$, $p = 0.0719$). However, orders under S were significantly greater than orders under C ($F(1, 67) = 16.80$, $p = 0.0001$), and orders under O were significantly greater than orders under C ($F(1, 67) = 35.75$, $p < 0.0001$). Table 7 shows that this same pattern of significant differences appears to be present in the average order, the round 1 order, and the follow-up question.

Though orders appear to be increasing over time in the high-profit condition, the effect was not significant (Wilks' lambda $= 0.641$, $F(19, 49) = 1.44$, $p = 0.1502$). We also found no significant interaction between payment scheme and round (Wilks' lambda $= 0.719$, $F(38, 98) = 0.46$, $p = 0.9957$). The actual mean demands were 10.787, 10.679, and 10.670, under O, S, and C, respectively.

Low-Profit Condition. For the low-profit condition, the repeated measures generalized linear model showed that payment scheme significantly affected ordering behavior ($F(2, 57) = 7.15$, $p = 0.0017$). We again found that average orders were highest under payment scheme O and lowest under payment scheme C. However, the differences that were significant were not the same as in the high-profit condition. Order quantities under O were significantly greater

Figure 3 **Average Order Quantities in Each Round in Study 3 for (a) High-Profit and (b) Low-Profit Conditions Under Schemes O, S, and C**



than orders under S ($F(1, 57) = 9.22$, $p = 0.0036$), but orders under S were not significantly greater than orders under C ($F(1, 57) = 0.19$, $p = 0.666$). Orders under O were significantly greater than orders under C ($F(1, 57) = 12.04$, $p = 0.001$). Again, Table 7 shows that these significance patterns appear in the average order, the round 1 order, and the follow-up question. Though from Figure 3 it appears that orders were decreasing over time, the effect was not significant (Wilks' lambda $= 0.728$, $F(19, 39) = 0.77$, $p = 0.73$). We also found no significant interaction between payment scheme and round (Wilks' lambda $= 0.440$, $F(38, 78) = 1.04$, $p = 0.428$). The actual mean demands were 10.523, 10.570, and 10.690, under O, S, and C, respectively.

Summary. Study 3 examines the effect of payment schemes for high- and low-profit conditions. We find that in both conditions, payment scheme significantly

Table 7 Mean and Standard Deviations of Order Quantities, and Significance Tests for Differences Between Payment Schemes in Study 3

Payment scheme	Mean order quantity (standard deviation in parentheses): Average over 20 rounds	Round 1	Follow-up question	N
	High-profit condition			
O	11.821 (1.336)	11.250 (2.691)	12.250 (1.595)	24
S	11.233 (1.020)	10.522 (1.344)	11.610 (1.373)	23
C	9.900 (0.892)	9.130 (1.359)	10.261 (0.964)	23
	Contrast tests			
$q^o - q^s$	0.588	0.728	0.640	
$q^s - q^c$	1.333***	1.392*	1.349**	
$q^o - q^c$	1.921***	2.120***	1.989***	
	Low-profit condition			
O	10.478 (1.270)	11.100 (2.900)	10.350 (0.988)	20
S	9.305 (0.974)	9.600 (2.210)	8.850 (1.089)	20
C	9.137 (1.382)	9.050 (2.012)	8.700 (1.719)	20
	Contrast tests			
$q^o - q^s$	1.173**	1.500	1.500***	
$q^s - q^c$	0.168	0.550	0.150	
$q^o - q^c$	1.340***	2.050**	1.650***	

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

affects ordering decisions. Study 3 also provides a robustness check of the prospective accounting model. In support of the prospective accounting model, we find that in both conditions orders exhibit a decreasing pattern $q^O > q^S > q^C$. Nevertheless, not all of these differences are significant. Specifically, for the high-profit condition, we find significant support for $q^S > q^C$ and $q^O > q^C$, but not for $q^O > q^S$. On the other hand, for the low-profit condition, we find significant support for $q^O > q^S$ and $q^O > q^C$, but not for $q^S > q^C$.

We offer the following explanation for this distortion. Because of the different profit parameters, the amount of order-time payments subject to underweighting under prospective accounting is different under schemes O, S, and C. Under the high-profit condition, the magnitudes of the order-time payments per unit under schemes O, S, and C are \$1, \$0, and \$3, respectively. Thus, prospective accounting has a much greater impact on payment scheme C compared to schemes O and S. This is consistent with our observations that differences between O and C and between S and C are significant, but the difference between O and S is not. Similarly, for the low-profit condition, the magnitudes of order-time payments per unit are \$3, \$0, and \$1 under O, S, and C, respectively. Thus, prospective accounting has a much greater impact on payment scheme O compared to schemes S and C, leading to significant differences between O and C and between O and S, but not between S and C.

Table 8 Contrasts Between Payment Schemes in the Average Observed Order Quantities in Study 3 and the Theoretical Order Quantities According to the Prospective Accounting Model

Payment scheme contrast	Predicted differences under prospective accounting model with $\beta = 0.7125$, $\alpha = 0.8137$: High profit	Low profit	Average differences observed in Study 3: High profit	Low profit
$q^O - q^S$	0.734	1.740	0.588	1.173
$q^S - q^C$	1.167	0.458	1.333	0.168

To further understand this phenomenon, we calculate the optimal order quantities based on the prospective accounting model using the estimated factors β and α obtained from Study 1 (specifically, the estimates from the NH method). The results are shown in Table 8, along with the observed differences in order quantities in Study 3. From this table, we see that the prospective accounting model is consistent with the findings of Study 3 in terms of the relative differences between order quantities under O, S, and C.

Structural Parameter Estimates. To obtain an estimate for the parameters β and α in the high- and low-profit conditions, we follow the same procedure as in Studies 1 and 2, but control for the pull-to-center effect. We assume the pull-to-center effect is of the same magnitude across each payment scheme and estimate it using maximum likelihood (the "TS" step in the two-step procedure proposed by Olivares et al. 2008). The results are reported in Table 9. These estimates again confirm our observations from the order quantity comparisons under schemes O, S, and C.

Expected Profits. For the high-profit condition, expected profits were significantly affected by payment scheme ($F(2, 67) = 11.98$, $p < 0.001$). The expected profits were highest under payment scheme O. Average expected per-round profits by condition were 26.690 (0.500), 26.552 (0.679), and 25.567 (1.222) for O, S, and C, respectively. Contrast tests show that all differences between conditions are significant differences at the $p < 0.05$ level except between O and S. In other words, participants under O and S performed significantly better than those under C. For the low-profit condition, expected profits were significantly affected by payment scheme ($F(2, 57) = 6.30$, $p = 0.003$). The average expected per-round profits were lowest under payment scheme O. Average expected profits by condition were 4.858 (1.555), 5.881 (0.743), and 5.847 (0.496) for O, S, and C, respectively. Contrast tests show that all differences between conditions are significant differences at the $p < 0.05$ level except between S

Table 9 Parameter Estimates for Study 3

Profit condition	Parameter	Model N1	Model N2	Model NH
High	$\hat{\beta}$	0.9638 (0.0303)	0.7649 (0.0353)**	0.8514 (0.0829)
High	$\hat{\alpha}$	0.7757 (0.0105)**	0.7848 (0.0115)**	0.7801 (0.0315)**
Low	$\hat{\beta}$	0.7959 (0.0138)**	0.8127 (0.0174)**	0.8038 (0.0459)**
Low	$\hat{\alpha}$	1.0845 (0.0352)	0.9294 (0.0390)*	1.0187 (0.1179)

Note. Standard errors are in parentheses.
$^{*}p < 0.05$; $^{**}p < 0.01$.

and C. In other words, participants under C and S performed significantly better than those under O. The expected profit analysis above also demonstrates how the effect of payment scheme interacts with the pull-to-center effect and the resulting effectiveness of ordering behavior. For the high-profit condition, the pull-to-center effect is mitigated by the effect of scheme O, but exacerbated by the effect of scheme C. Conversely, for the low-profit condition, the pull-to-center effect is mitigated by the effect of scheme C, but exacerbated by the effect of scheme O.

On the Asymmetry of the Pull-to-Center Effect. Schweitzer and Cachon (2000) and Bolton and Katok (2008) found that the pull-to-center effect is stronger, i.e., the deviation from the optimal order quantity toward the center of the demand distribution is larger, in the low-profit condition. They suggest stockout aversion or that the high-profit condition is more "intuitive" as possible explanations, but do not provide substantive evidence. The framing of the newsvendor problem in their papers is similar to the wholesale price contract (i.e., scheme O). Thus, to examine this issue, we compared the level of deviation from the optimal order quantity under high- and low-profit conditions for the three payment schemes.

We found that under payment scheme O, participants deviated farther from the optimal solution (toward the center of the distribution) under the low-profit condition compared to the high-profit condition ($F(5, 124) = 13.70$, $p = 0.0003$). Average orders under the low-profit condition were 2.478 above the optimal order 8, but average orders under the high-profit condition were only 1.179 below the optimal 13. This is the same asymmetry effect observed by Schweitzer and Cachon (2000) and Bolton and Katok (2008).

However, under payment scheme S, participants deviated approximately the same distance from the optimal solution in both the high- and low-profit conditions, and the deviations were not significantly different ($F(5, 124) = 1.70$, $p = 0.194$). Average orders under the low-profit condition were 1.305 above the optimal order 8, and average orders under the high-profit condition were 1.767 below the optimal 13. Thus, we observed that the asymmetry effect disappears under payment scheme S.

Finally, under payment scheme C, participants deviated farther from the optimal solution (toward the center of the distribution) under the high-profit condition compared to the low-profit condition ($F(5, 124) = 30.70$, $p < 0.0001$). Average orders under the high-profit condition were 3.100 below the optimal 13, but average orders under the low-profit condition were only 1.138 above the optimal order 8. Thus, the asymmetry effect in this case is in the opposite direction to that of payment scheme O.

5. Discussion

In this paper, we find that payment schemes have a significant behavioral effect on ordering decisions in the newsvendor problem. Our findings help us gain insights into how human subjects account for payments. We provide evidence that order-time payments receive less weight than the demand-time payments, which is consistent with the prospective accounting model. We also demonstrate that the framing of the payments is sufficient to induce differences in ordering behavior, even if the actual timing of payments are the same across payment schemes, and that the prospective accounting effect is robust for high- and low-profit conditions.

Our laboratory findings are also related to other behavioral theories. Schweitzer and Cachon (2000) applied the idea of "recency" to obtain a "chasing demand" heuristic, in which individuals use their previous order quantity as an anchor and adjust toward the previous demand realization. One could apply the same idea of recency to obtain a heuristic in which individuals anchor on their previous order quantity and adjust by putting a greater weight on the most recent payment feedback (the payments after demand realization). This would also yield behavior consistent with a decision maker that underweights the order-time payments, which is similar to what Gourville and Soman (1998) called "payment depreciation." (Payment depreciation is essentially the second half of the prospective accounting rule that states that events are weakly coupled looking backward in time.) However, such an adjustment heuristic does not explain why we find differences in ordering behavior between payment schemes in the first round of order decisions in our experiments. Another behavioral effect that is relevant to our setting is "debt aversion." Because the newsvendor is in debt under payment scheme S, debt aversion would predict larger orders under scheme O than scheme S. However, it does not explain why payment scheme O leads to orders larger than the expected-profit-maximizing solution in Study 1. Furthermore, Prelec and Loewenstein (1998, p. 14) suggested that "debt aversion is a general implication of prospective accounting," which implies that debt aversion is a result of mental accounting rather than a separate behavioral effect. Study 2 demonstrates that indeed our observations are not solely driven by the actual debt position over time, but how the individual mentally processes the payments over time.

The behavioral effect of payment schemes on orders has direct implications on the newsvendor's expected profit, the supplier's revenue, and the customer's service level. Therefore, from the newsvendor's standpoint, one should strategically select the appropriate payment scheme to encourage a most efficient decision outcome. In our experiments, payment scheme S, which conducts all payments after the demand realization, encourages equal weighting of all payments and achieves the optimal order when the overage and underage costs are equal. However, payment schemes can also be used to mitigate other behavioral biases. For example, for the pull-to-center effect, the payment scheme that would lead to order quantities closest to the optimal solution is the payment scheme O under the high-profit condition and the payment scheme C under the low-profit condition. Furthermore, Study 2 demonstrates that even if the actual payment contract does not have a timing of the payments that induces an optimal ordering decision, one can simply rewrite the framing of payment scheme to encourage optimal behavior. On the other hand, the supplier and the customer would like to choose a payment scheme that induces the highest order, which, according to our experiments, is the wholesale price scheme (payment scheme O). In addition to practitioners, our results may inform future newsvendor experiments, as individuals weigh the payments correctly only under the framing of payment scheme S.

Our results also help us gain insight on the behavioral effect of financial contracts in practice. For instance, suppliers often offer retailers trade credit, allowing retailers to delay payment for goods until they make the sale, hoping that this will encourage higher orders. This intended effect of trade credit is captured by the time-discounting model in §3.3. When capital constraint are not an issue and the interest rate is negligible, the practice of trade credit (corresponding to the payment scheme S) might inadvertently lower the retailer's order quantity relative to that without trade credit (corresponding to the payment scheme O), as shown by our experiments. When the interest rate is significant or the capital constraint is binding, however, the prospective accounting effect may become second order. Similarly, though we show that providing the newsvendor revenue-based loans may decrease the inventory order, such a behavioral effect may also be dominated by the time-discounting effect due to the tangible interest rate benefit. An interesting direction would be to determine the relative magnitudes of these opposing effects in practice and estimate the impact of mental accounting empirically.

Another application is in supply chain contract design and coordination (see Cachon 2003 for a review). A wholesale price contract typically has payment transactions resembling the payment scheme O in this paper. Our results suggest that the retailer may place larger-than-optimal orders due to prospective accounting, reducing some of the supply chain inefficiency due to double marginalization (Lariviere and Porteus 2001). If the supplier can estimate the retailer's underweighting factor as we did in §4, then she may coordinate the supply chain by setting the wholesale price equal to the unit production cost divided by the underweighting factor. Under a buyback contract, the retailer receives a refund for leftover inventory after the demand realization. To the retailer, the refund payment is likely to be weighted more than the purchase cost incurred at the order time. Thus, the supplier may exploit this effect to achieve supply chain coordination by offering a smaller buyback price for leftover inventory. For a similar reason, under a revenue-sharing contract, with prospective accounting, the supplier may be able to charge a higher wholesale price to the retailer and still achieve supply chain coordination. It would be interesting to empirically investigate these potential implications, although we acknowledge that there are many additional factors working simultaneously in real-world contract settings (see Zhang et al. 2012).

Finally, several extensions to our study merit further research. First, it would be interesting to test how individuals react to a payment scheme switch in the newsvendor problem. This could enable the comparison of each individual's decisions over different payment schemes. If this is to be carried out, we caution that some extra care should be taken to control for the recency and learning effects across the scheme switch. Second, it would also be interesting to test how individuals place inventory orders among multiple suppliers who offer different payment schemes. This could further shed light on the effect of payment schemes in a competitive environment.

Acknowledgments
The authors thank Rick Larrick, Jack Soll, Jeannette Song, Daniel Feiler, and Gérard Cachon for helpful discussions on this paper, and research assistants Nathan Glencer and Lucy Zheng for their help with running the experiments.

Appendix. Comparison to Prelec and Loewenstein (1998) Model
Prelec and Loewenstein (1998) incorporated prospective accounting through a "double-entry" model. In the context of the newsvendor problem, the "double-entry" feature of their consumer model suggests that the newsvendor imputes payments twice: from the vantage points of the order and the demand realization. Their model is quite sophisticated—including loss aversion, time discounting, etc.—but we focus only on the coupling feature. In the newsvendor context, "coupling" qualifies the prospective accounting rule by allowing only partial appreciation of payments looking forward. Thus, we add a coupling term b, $0 \le b \le 1$, to denote how strongly an outgoing order-time payment is "buffered" by the thought of incoming demand-time payments. Similarly, we add a coupling term a, $0 \le a \le 1$, to denote how strongly an incoming order-time payment is "attenuated" by the thought of outgoing demand-time payments. The resulting reward functions are

$$R^i(q, D) = \begin{cases} -cq + bp\min(q, D) + p\min(q, D) & \text{if } i = O, \\ 2[(p-c)\min(q, D) - c\max(q-D, 0)] & \text{if } i = S, \\ (p-c)q - ap\max(q-D, 0) - p\max(q-D, 0) & \text{if } i = C. \end{cases}$$

The above formulation also results in the prediction $q^O > q^S > q^C$. However, for simplicity, our prospective accounting model uses a simpler formulation that merely underweights the order-time payments. Under scheme O, the cost is "buffered," so we use the notation $\beta < 1$. Under scheme C, the net profit is "attenuated," so we use the notation $\alpha < 1$.

References
Babich V, Sobel MJ (2004) Pre-IPO operational and financial decisions. *Management Sci.* 50(7):935–948.

Bendoly E, Donohue K, Schultz K (2006) Behavioral operations management: Assessing recent findings and revisiting old assumptions. *J. Oper. Management* 24(6):737–752.

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

Bolton GE, Ockenfels A, Thonemann U (2010) Better decisions? Why experience managers aren't always right. *Internat. Commerce Rev.* 9(1):49–57.

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

Burgstahler D, Dichev I (1997) Earnings management to avoid earnings decreases and losses. *J. Accounting Econom.* 24:99–126.

Buzacott JA, Zhang RQ (2004) Inventory management with asset-based financing. *Management Sci.* 50(9):1274–1292.

Cachon GP (2003) Supply chain coordination with contracts. de Kok AG, Graves SC, eds. *Handbooks in Operations Research and Management Science: Supply Chain Management* (North-Holland, Amsterdam), 229–332.

Croson R, Donohue K (2005) Upstream versus downstream information and its impact on the bullwhip effect. *System Dynam. Rev.* 21(3):249–260.

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

Feiler DC, Tong JD, Larrick RP (2012) Biased judgment in censored environments. *Management Sci.*, ePub ahead of print December 19, http://dx.doi.org/10.1287/mnsc.1120.1612.

Fisher M, Raman A (1996) Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* 44(1):87–99.

Frederick S, Loewenstein G, O'Donoghue T (2002) Time discounting and time preference: A critical review. *J. Econom. Literature* 40(2):351–401.

Gourville JT, Soman D (1998) Payment depreciation: The behavioral effects of temporally separating payments from consumption. *J. Consumer Res.* 25(2):160–174.

Gupta D, Wang L (2009) A stochastic inventory model with trade credit. *Manufacturing Service Oper. Management* 11(1):4–18.

Haley CW, Higgins RC (1973) Inventory policy and trade credit financing. *Management Sci.* 20(4):464–471.

Heath C, Soll J (1996) Mental budgeting and consumer decisions. *J. Consumer Res.* 23(1):40–52.

Ho T-H, Zhang J (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, Lim N, Cui TT (2010) Reference dependence in multilocation newsvendor models: A structural analysis. *Management Sci.* 56(11):1891–1910.

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

Katok E, Wu DY (2009) Contracting in supply chains: A laboratory investigation. *Management Sci.* 55(12):1953–1968.

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

Laibson D (1997) Golden eggs and hyperbolic discounting. *Quart. J. Econ.* 112(2):443–477.

Lariviere MA, Porteus EL (2001) Selling to the newsvendor: An analysis of price-only contracts. *Manufacturing Service Oper. Management* 3(4):293–305.

Loch CH, Wu Y (2008) Social preferences and supply chain performance: An experimental study. *Management Sci.* 54(11): 1835–1849.

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

Ng CK, Smith JK, Smith RL (1999) Evidence on the determinants of credit terms used in interfirm trade. *J. Finance* 54(3):1109–1129.

Olivares M, Terwiesch C, Cassorla L (2008) Structural estimation of the newsvendor model: An application to reserving operating room time. *Management Sci.* 54(1):41–55.

O'Sullivan K (2007) Financing the chain. *CFO Magazine* (February 1), http://cfo.com/article.cfm/8581992.

Peterson MA, Rajan RG (1997) Trade credit: Theories and evidence. *Rev. Financial Stud.* 10(3):661–691.

Prelec D, Loewenstein G (1998) The red and the black: Mental accounting of savings and debt. *Marketing Sci.* 17(1):4–28.

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

Shafir E, Thaler RH (2006) Invest now, drink later, spend never: On the mental accounting of delayed consumption. *J. Econom. Psych.* 27(5):694–712.

Shefrin H, Statman M (1985) The disposition to sell winners too early and ride losers too long: Theory and evidence. *J. Finance* 40(3):777–790.

Song J-S, Tong J (2012) Payment schemes, financing costs, and inventory management. Working paper, Duke University, Durham, NC.

Sterman JD (1989) Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Sci.* 35(3):321–339.

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

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

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

Thaler RH (1999) Mental accounting matters. *J. Behav. Decision Making* 12(3):183–206.

Tversky D, Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* 211(4481):453–458.

Xu X, Birge JR (2004) Joint production and financing decisions: Modeling and analysis. Working paper, University of Chicago, Chicago.

Xu Y, Zhang J (2010) On the selection of supply chain coordinating contracts. Working paper, University of Maryland, College Park.

Zhang Y, Donohue K, and Cui TH (2012) Contract preferences for the loss averse supplier: Buyback versus revenue sharing. Working paper, University of Minnesota, Minneapolis.