

# Contracting in Supply Chains: A Laboratory Investigation

Elena Katok

Smeal College of Business, Pennsylvania State University, University Park, Pennsylvania 16802, ekatok@psu.edu

Diana Yan Wu

School of Business, University of Kansas, Lawrence, Kansas 66045, dianawu@ku.edu

The coordination of supply chains by means of contracting mechanisms has been extensively explored theoretically but not tested empirically. We investigate the performance of three commonly studied supply chain contracting mechanisms: the wholesale price contract, the buyback contract, and the revenue-sharing contract. The simplified setting we consider utilizes a two-echelon supply chain in which the retailer faces the newsvendor problem, the supplier has no capacity constraints, and delivery occurs instantaneously. We compare the three mechanisms in a laboratory setting using a novel design that fully controls for strategic interactions between the retailer and the supplier. Results indicate that although the buyback and revenue-sharing contracts improve supply chain efficiency relative to the wholesale price contract, the improvement is smaller than the theory predicts. We also find that although the buyback and revenue-sharing contracts are mathematically equivalent, they do not generally result in equivalent supply chain performance.

*Key words:* supply chain contracts; experimental economics

*History:* Received June 12, 2006; accepted July 28, 2009, by Aleda Roth, operations and supply chain management. Published online in *Articles in Advance* October 16, 2009.

## 1. Introduction and Motivation

Previous studies have shown a great deal of interest in the analysis of contracting mechanisms that can be used to coordinate supply chains. The alignment of the economic incentives of supply chain partners is important because a supply chain that consists of individual firms, each interested in its own welfare, yields decentralized decisions that are usually inefficient. Much of the past research has focused on the analytical design of contracting arrangements to eliminate this inefficiency (see Cachon 2003 for a review).

In this study we undertake an experimental analysis of the simplest supply chain contracting setting that has been analyzed theoretically, in which the retailer faces the classic newsvendor problem and orders from a supplier, the supplier has no capacity constraint, and delivery occurs instantaneously. This simple model is a building block for much of the contracting literature in operations management, and as such it represents a logical starting point for our initial laboratory study. In this setting, whenever a supplier charges a wholesale price in excess of his own production cost, double marginalization (Spengler 1950) causes the retailer to order less than the channel-optimal amount. These smaller retailer orders imply that all channel partners forgo potential profits; consequently, methods for avoiding this inefficiency are valuable because they can make both parties better off. To coordinate the

supply chain, a contract must give the retailer the incentive to order the same amount that would be optimal in a centralized setting. Cachon (2003) provides a review of the analytical work that investigates various contractual arrangements that facilitate this coordination.

One useful class of arrangements shares the demand risk between the retailer and the supplier by making the supplier's profit depend on realized sales. Cachon and Lariviere (2005) look at two such risk-sharing contracts, the buyback contract and the revenue-sharing contract, and show that the two are mathematically equivalent in the strongest possible way, meaning that they generate identical outcomes for both players for each possible realization of the stochastic demand. In the buyback contract, the supplier pays the retailer a rebate on all unsold units, thus assuming some of the risk associated with overordering. In the revenue-sharing contract, the supplier induces a higher retailer order through a lower wholesale price, but in return he receives a portion of the gross revenue. Buyback contracts are quite common in industries such as publishing, computer software and hardware, and pharmaceuticals (Padmanabhan and Png 1995), whereas revenue-sharing contracts are observed in the video-rental industry (Cachon and Lariviere 2005).

These two types of contracts and the simple analytical model upon which they are based have generated significant scholarly interest and follow-up studies. The problem has been extended to two stages (Donohue 2000) and to secondary markets (Rudi et al. 2001, Lee and Whang 1999). Other extensions consider service levels and stockouts (Choi et al. 2004), flexibility (Kamrad and Siddique 2004), incomplete information (Corbett et al. 2004), procurement contracts (Wu and Kleindorfer 2005), option contracts (Burnetas and Ritchken 2005, Kleindorfer and Wu 2003), warranty contracts (Balachandran and Radhakrishnan 2005), and target-rebate contracts on false failure returns (Ferguson et al. 2006).

These, along with most other analytical models of supply chain coordination, assume that contracting parties are fully rational expected-profit maximizers. But there is now a growing body of evidence based on laboratory experiments (Schweitzer and Cachon 2000, Bolton and Katok 2008) as well as some field studies (Corbett and Fransoo 2007) showing that retailers have difficulty making optimal decisions even in the simplest settings. Therefore, it remains unclear whether the theoretical gains from more complex contractual arrangements are likely to be achieved in practice.

We extend prior research on supply chain coordination and contracting in three ways. First, we investigate retailers' behavior when faced with coordinating contracts and find those contracts to be significantly less effective than the theory suggests. Second, we study the suppliers' behavior in structuring coordinating contracts and find that suppliers do not offer contracts that fully coordinate the supply chain, even when retailers are programmed to order optimally, given contract parameters. Third, we compare two equivalent coordinating contracts—the buyback contract and the revenue-sharing contract—from both the retailers' and the suppliers' perspectives and find that although they do not always induce equivalent outcomes, most of the differences disappear with experience. Loss aversion (Kahneman and Tversky 1979), which Ho and Zhang (2008) have already shown to be important in contracting, can explain these initial differences. We use a controlled laboratory setting designed to conform to the assumptions of the contracting models we are testing. Our design is novel in that it eliminates interpersonal interactions, and thus it allows us to attribute any deviations from theoretical benchmarks to individual decision-making biases.

This paper is structured as follows. In the next section, we position our work relative to existing analytical, empirical, and experimental research. In §3, we describe the details of the experimental design and the laboratory protocol. We develop our experimental hypotheses in §4, and in §5, we report the

results. In §6, we summarize our findings, identify the limitations of our study and directions for future research, and discuss the managerial implications of our results.

## 2. Analytical Background and Related Literature

### 2.1. Analytical Background

In the baseline model (Spengler 1950), the wholesale price contract, the retailer orders  $q$  units from a supplier at a wholesale price of  $w$  per unit. The retailer faces an exogenous stochastic demand with cumulative distribution  $F(\cdot)$  and an exogenous market price  $p$ , and he suffers losses whenever his actual order quantity  $q$  differs from the realized demand  $D$ . The retailer maximizes his expected profit by balancing the cost of ordering too much or too little, and to do that he sets his order  $q$  to satisfy

$$F(q) = \frac{p - w}{p},$$

which is known as the *critical fractile*.

In contrast, the supplier faced with the wholesale price contract incurs no risk, because when the production cost is  $c$  and the wholesale price is  $w > c$ , he simply makes a profit of  $(w - c) \times q$  on the retailer's entire order. The wholesale price that maximizes the supplier's expected profit in the wholesale price contract depends on the demand distribution. If  $F(\cdot)$  is uniform from  $A$  to  $B$ , then the optimal wholesale price  $w^*$  is given by

$$w^* = \min \left\{ p, \frac{c}{2} + \frac{p}{2} \frac{B}{B - A} \right\}. \quad (1)$$

If we let  $0 < \lambda < 1$  be the retailer's share of the total profit, a continuum of coordinating risk-sharing contracts can be constructed, one for each  $\lambda$ . If the supplier uses a buyback contract in which he charges the retailer  $w_{BB}$  per unit, and then refunds  $b$  per unit for all units unsold at the end of the selling season, such a contract coordinates the supply chain (i.e., induces the retailer to place the channel-optimal order) when pairs of parameters  $\{w_{BB}, b\}$  satisfy

$$\begin{aligned} b &= (1 - \lambda)p, \\ w_{BB} &= b + \lambda c. \end{aligned} \quad (2)$$

Cachon and Lariviere (2005) show that the revenue-sharing contract in which the retailer pays  $w_{RS}$  per unit ordered and an additional  $r$  per unit sold is equivalent to the buyback contract  $\{w_{BB}, b\}$  when the cost for units sold and unsold is the same under both arrangements:

$$\begin{aligned} \text{Per-unit cost of units sold: } & w_{BB} = w_{RS} + r, \\ \text{Per-unit cost of units unsold: } & w_{BB} - b = w_{RS}. \end{aligned} \quad (3)$$

For a uniformly distributed demand,  $D \sim U(A, B)$ , expected sales given an order of  $q$  are given by

$$E[S] = \left(\frac{A+q}{2}\right) \left(\frac{q-A}{B-A}\right) + q \left(\frac{B-q}{B-A}\right),$$

and correspondingly, the expected profit amounts for the retailer, the supplier, and the total supply chain are given by

$$\begin{aligned} \pi^R &\equiv E[\pi^{\text{Retailer}}] = (p-r)E[S] - wq + b(q-E[S]), \\ \pi^S &\equiv E[\pi^{\text{Supplier}}] = (w-c)q + rE[S] - b(q-E[S]), \\ \pi^T &\equiv E[\pi^{\text{Total}}] = pE[S] - cq. \end{aligned} \quad (4)$$

The retailer's expected share of the total profit,  $\lambda$ , is then  $\lambda = \pi^R / \pi^T$ .

## 2.2. Laboratory Research

Laboratory studies that investigate the retailer's contracting behavior have focused almost exclusively on wholesale price contracts. Schweitzer and Cachon (2000) find that retailers place orders that tend to be between the optimal orders and the average demand, and note that this "pull-to-center" effect cannot be explained by risk preferences, loss aversion, or prospect theory. They suggest that this behavior is consistent with the minimization of ex post inventory error and the anchoring and insufficient adjustment heuristic. Bolton and Katok (2008) also find the pull-to-center effect and additionally show that performance improves over time with extensive experience, although slowly, and that requiring decision makers to place standing orders<sup>1</sup> speeds up learning substantially. Lurie and Swaminathan (2009) report a similar finding, specifically that feedback that is too frequent can degrade performance and slow down learning. Benzion et al. (2008) vary the demand distribution and find that while orders are affected by both the average demand and the last-period demand, this bias is weakened slowly over time. This implies that participants learn over time not to chase demand. Bostian et al. (2008) find that an adaptive learning model explains the pull-to-center effect. Overall, the evidence that human players fail to place expected profit-maximizing orders when faced with the newsvendor problem is fairly conclusive.

Several recent papers study suppliers' and retailers' behavior jointly and find that profits tend to be distributed more equitably and the efficiency of coordinating contracts is lower than the standard theory predicts. Keser and Paleologo (2004) find that in a stochastic demand setting, newsvendor retailers are

likely to reject contracts with high wholesale prices, and therefore suppliers tend to choose wholesale price contracts that split profits approximately equally when the entire order is sold. Ho and Zhang (2008) look at a bilateral monopoly setting (deterministic downward-sloping linear demand) and report that two-part tariffs and quantity discount contracts fail to coordinate the supply chain or even achieve a level of efficiency that is significantly above the wholesale price contract efficiency levels. They attribute coordination failure to a combination of loss aversion and bounded rationality. Loch and Wu (2008) also study the wholesale price contract in a bilateral monopoly setting, but in their study the retailer and the supplier interact repeatedly. They find that efficiency decreases when players are concerned about status, and increases when they are concerned about their relationship. The present study is the first laboratory investigation of the performance of risk-sharing coordinating contracts.

Standard operations management models follow economic assumptions about human behavior, including that players are fully rational expected-profit maximizers. This assumption implies that players want to maximize only their own expected profit and have the cognitive ability to do so. In practice, human decision makers negotiate contracts, and they may violate this standard theoretical assumption for one of the following reasons:

(1) *Bounded rationality*—Decision makers want to maximize their expected profit, but make errors in doing so, or resort to heuristics.

(2) *Different utility functions*—Decision makers maximize a utility function that includes other attributes in addition to the expected profit, such as loss aversion or concern for fairness.

When two human players interact, it is well established that they are motivated by concerns for fairness. Cui et al. (2007) incorporate these fairness concerns into an analytical model of contracting. De Bruyn and Bolton (2008) report on a metastudy that shows that a simple model incorporating fairness explains a large variety of data. Loch and Wu (2008) provide a review of the way that general social preferences affect operations management models.

But to understand the effect of social preferences, it is important to recognize which deviations from theoretical benchmarks are due to social preferences, and which are due to other decision-making biases. We designed our study with this initial step in mind. In the present study, human retailers deal with computerized suppliers, and human suppliers deal with computerized retailers. All human decision makers in our study interact with computerized players programmed to act according to theory. In Figure 1 we depict the scope of the contracting problem and show how our study fits into the larger picture. We call

<sup>1</sup> In this setting, a *standing order* refers to a restriction that forces a retailer to place one order that is used for several consecutive periods (in Bolton and Katok 2008, this was 10 periods).

Figure 1 Scope of the Contracting Problem and Our Study

		Supplier is...	
		Not human	Human
Retailer is...	Not human	<i>Standard theory:</i> Full rationality Expected-profit maximization	<i>The supplier game:</i> Bounded rationality Individual biases <b>Decision:</b> $\{wblr\}$
	Human	<i>The retailer game:</i> Bounded rationality Individual biases <b>Decision:</b> $\{q\}$	<i>Practice:</i> Bounded rationality Individual biases Social preferences <b>Decision:</b> $\{q\}$ and $\{wblr\}$

the setting with human retailers interacting with computerized suppliers the *retailer game* and the setting with human suppliers interacting with computerized retailers the *supplier game*. In both games, the retailer faces a single-period newsvendor problem.

Our study design eliminates the possibility that players are motivated by social preferences such as fairness, and this is both a strength and a weakness. It is a strength because any deviations from theory that we discover must be attributed either to individual biases or to bounded rationality, but definitely not to social preferences.<sup>2</sup> Bounded rationality, and to some extent individual biases, can be eliminated or at least lessened through education and the use of decision-support tools, but social preferences cannot be eliminated (nor is it desirable to try to do so). A thorough understanding of which deviations from theory can be overcome through the use of technology, and which instead should be included in models to make the models more valid, is valuable. A study that looks at how two human decision makers negotiate contracts is an important next step, and in the concluding section we discuss this direction for future research, informed by the present study.

### 3. Experimental Implementation

#### 3.1. Experimental Design

Contracting arrangements in our experiments are the wholesale price contract ( $W$ ), the buyback contract ( $BB$ ), and the revenue-sharing contract ( $RS$ ). In all treatments we set the supplier's production cost to be  $c = 3$  and the retail price to be  $p = 12$  to create a setting with potential high supply chain profits. We focus on the high-profit condition (critical fractile  $> 1/2$ ) in this paper because it is a setting with greater possible gains from coordination.

To investigate the effect of loss aversion on behavior, we use three different uniform customer demand

conditions. In the *DLOW* condition,  $D \sim U(0, 100)$ , demand is potentially low (as low as 0), and both retailers and suppliers can lose money under optimal coordinating contracts. In the *DHIGH* condition,  $D \sim U(50, 150)$ , demand is potentially greater, and retailers can lose money under optimal coordinating contracts, but suppliers cannot. Under the wholesale price contract, by contrast, suppliers cannot lose money with either of the demand distributions, and retailers can lose with both.

We call the third demand condition *DHIGH with DLOW decision frame* (DHIGH/LOW). The actual demand distribution is the same as DHIGH, but we describe it in a different way: as 50 guaranteed units and an additional number of units from 0 to 100, so that  $D = 50 + X$ , where  $X \sim U(0, 100)$ .  $X$  thus follows the same distribution as the demand in the DLOW condition. Participants are asked to decide on the number of units to order in addition to the 50 guaranteed units, so the order quantity is from 0 to 100, as in the DLOW condition. In this demand condition, suppliers cannot earn a negative profit, but if  $q$  is high and  $X$  is low, suppliers can lose money relative to the revenue they are guaranteed from selling the 50 units ( $50 \times [w + r - c]$ ). Thus, the DHIGH/LOW demand condition induces loss aversion through the manipulation of framing. This allows us to explore loss aversion as a potential explanation for any differences we observe between the buyback and revenue-sharing contracts.

The final factor we manipulate in this study is experience. To test for the effect of experience, we conduct each treatment twice, first with inexperienced participants, and then again with participants who had prior experience in the same role (either as retailer or supplier) with a different contract. Each session included 100 rounds, so each participant played for 200 rounds in total: 100 rounds in the inexperienced session, followed by 100 rounds in the experienced session.

In summary, our study manipulates the decision maker's role (retailer game and supplier game), the customer demand distribution (DLOW, DHIGH, and DHIGH/LOW), and experience (inexperienced and experienced). We summarize all treatments and sample sizes in Table 1. In total, 200 subjects participated in our study.

#### 3.2. Experimental Protocol

All experimental sessions followed the same protocol. Participants arrived at the computer lab at a pre-specified time and read experimental instructions that describe the rules of the game, the use of the software, and the payment procedures (see the online appendix, provided in the e-companion).<sup>3</sup> After all

<sup>2</sup> We thank an anonymous referee for pointing out that the previous statement assumes that people are capable of consciously making the distinction, and that people generally change their behavior in this task based on the presence of human beings on the other end. In Katok and Pavlov (2009) we show that this is in fact the case in a bilateral monopoly setting.

<sup>3</sup> An electronic companion to this paper is available as part of the online version that can be found at <http://mansci.journal.informs.org/>.

**Table 1** Experimental Design and Sample Sizes

Contract	Retailer game				Supplier game			
	Inexperienced		Experienced		Inexperienced		Experienced	
	Demand	<i>n</i>	Demand	<i>n</i>	Demand	<i>n</i>	Demand	<i>n</i>
Wholesale price	DLOW	20	DLOW	20	DLOW	20	DLOW	20
	DHIGH	20	DHIGH	20	DHIGH	20	DHIGH	20
Buyback	DLOW	10	DLOW	10	DLOW	10	DLOW	10
	DHIGH	10	DHIGH	10	DHIGH	10	DHIGH	10
	DHIGH/LOW	10	DHIGH/LOW	10	DHIGH/LOW	10	DHIGH/LOW	10
Revenue-sharing	DLOW	10	DLOW	10	DLOW	10	DLOW	10
	DHIGH	10	DHIGH	10	DHIGH	10	DHIGH	10
	DHIGH/LOW	10	DHIGH/LOW	10	DHIGH/LOW	10	DHIGH/LOW	10

Note. Demand distributions: DLOW,  $D \sim U(0, 100)$ ; DHIGH,  $D \sim U(50, 150)$ ; DHIGH/LOW,  $D = 50 + X$ ,  $X \sim U(0, 100)$ .

participants had a chance to read the instructions, the experimenter read instructions to them aloud to ensure common knowledge, used PowerPoint slides to illustrate examples and formulas, and answered questions. Participants then completed 100 rounds under the first of the two contracts. After all participants finished this initial (“inexperienced”) session, we handed them additional instructions describing the contract used in the second (“experienced”) session, gave them a chance to read these new instructions, read the instructions to them aloud, and answered questions before the beginning of the second session. After completing the second session, participants were paid their actual earnings accumulated from both sessions, privately and in cash. Participants were not allowed to communicate during the experiment.

All sessions were conducted at the Laboratory for Economic Management and Auctions at the Smeal College of Business, Penn State University. Each session lasted approximately 75 minutes, and the average earnings, including a \$5 participation fee, were \$19. Participants were Penn State students recruited through a Web-based recruitment system, with cash being the only incentive offered. The majority of our participants were undergraduates from a variety of majors (77%), and the rest were graduate students. We compared the average earnings by student level, major, and gender for each session using a *t*-test and found no response biases by those demographic characteristics.<sup>4</sup>

### 3.3. Experimental Implementation

In the retailer game, the human decision maker plays the role of the retailer, and we set the wholesale

price  $w$  optimally for both demand conditions using Equation (1) as follows:

$$w_w^{\text{DLOW}} = \min \left\{ 12, \frac{3}{2} + \frac{12}{2} \frac{100}{100-0} \right\} = 7.5, \tag{5}$$

$$w_w^{\text{DHIGH}} = \min \left\{ 12, \frac{3}{2} + \frac{12}{2} \frac{150}{150-50} \right\} = 10.5.$$

For the buyback contract, we used Equation (2) and set  $\lambda = 1/3$  so that both parties could benefit from coordination to obtain  $w_{BB} = 9$  and  $b = 8$ .<sup>5</sup> We then constructed the equivalent revenue-sharing contract using Equation (3), with  $r = 9$  and  $w_{RS} = 1$ .

In the supplier game, the human decision maker plays the role of the supplier. Here, our aim was to better understand the extent to which suppliers are able and willing to offer contracts that coordinate. The design includes an automated retailer programmed to act in accordance with theory: the retailer is programmed to place expected-profit maximizing orders, given the contract offered by the human supplier. This feature eliminates strategic interactions and also provides consistent feedback. Because  $F(\cdot)$  is  $U(A, B)$ , the best-response order quantity of the automated retailer is given by

$$q^* = A + (B - A) \left( \frac{p - w - r}{p - b - r} \right), \tag{6}$$

where  $A = 0$  and  $B = 100$  in the DLOW condition and  $A = 50$  and  $B = 150$  in the DHIGH and DHIGH/LOW conditions. Given these parameters, the order quantity that maximizes the retailer’s expected profit under the wholesale price contract is 37.5 in the DLOW condition and 62.5 in the DHIGH condition. Under the two coordinating contracts, the optimal order quantity is 75 in the DLOW condition and 125 in the DHIGH and DHIGH/LOW conditions.

<sup>5</sup> We wanted to avoid  $\lambda = 1/2$  so as not to confound our results by a 50/50 split.

<sup>4</sup> The human subject approval for this study required us to store the decision data using subject IDs, and therefore we do not have a way of connecting these IDs to individuals. However, our recruitment system gives us a way to track individual participants and their earnings by session, and we used this information for the analysis.

We set  $b = r = 0$  for the wholesale price contract, and participants selected  $w$  only. In the buyback contract we set  $r = 0$ , and participants selected  $w$  and  $b$  simultaneously; in the revenue-sharing contract we set  $b = 0$ , and participants selected  $w$  and  $r$  simultaneously.

Because the retailer in the supplier game is automated, the system provides feedback to suppliers about the order quantity that will follow a proposed contract, specifically,  $q^*$  as defined by Equation (6). Each participant was allowed to try different contract parameters and observe the expected (but not the actual) outcome as many times as desired before making the final decision for the round. We repeated this procedure for all treatments in the supplier game ( $W$ ,  $BB$ , and  $RS$ ) to make certain our participants had access to the relevant information that the theory implicitly assumes they have, thus giving the theory its best shot.

#### 4. Research Hypotheses

Recall that the human decision makers make different decisions in the two games: in the retailer game the decision is the order quantity  $q$ , whereas in the supplier game the decision is a set of contract parameters ( $w, b, r$ ). However, the retailer's order quantity  $q$  can be a unifying metric for both games, because the total channel profit is always proportional to  $q$ . Therefore, we formulate our first three research hypotheses and conduct the corresponding data analysis in terms of  $q$ .

All four hypotheses apply to both the retailer game and the supplier game. (When we talk about the retailer's average order in the context of the supplier game, we mean the average order of the automated retailer that the human supplier induces.) The first hypothesis follows directly from the quantitative predictions of the standard theory.

**HYPOTHESIS 1 (THEORETICAL BENCHMARKS).** *The retailers' average orders for wholesale price contracts will be 37.5 in the DLOW condition and 62.5 in the DHIGH condition. The average orders for the buyback and revenue-sharing contracts will be 75 in the DLOW condition and 125 in the DHIGH and DHIGH/LOW conditions.*

Even if the data deviate from these precise theoretical benchmarks, the theory can still make useful qualitative predictions. The main point of the standard theory summarized in §2.1 is that the buyback and revenue-sharing contracts can induce higher orders than the wholesale price contracts. Our second hypothesis reflects this qualitative prediction.

**HYPOTHESIS 2 (COORDINATION).** *The retailers' average orders for the buyback and revenue-sharing contracts will be higher than for the wholesale price contract.*

Our third hypothesis links theoretical predictions with known behavioral biases. Previous studies have

documented a "pull-to-center" effect (Schweitzer and Cachon 2000, Bostian et al. 2008) in which average orders are located between the optimal orders and the average demand.<sup>6</sup> Schweitzer and Cachon (2000) note that the data, although inconsistent with many established behavioral models (risk preferences, prospect theory, loss aversion) are in fact consistent with (i) the "anchoring and insufficient adjustment" heuristic, and (ii) a preference for minimizing ex post inventory error. Although these two explanations can each account for the pull-to-center effect, they involve different adjustment patterns. The anchoring and adjustment heuristic implies that orders start close to the average demand and adjust over time in the direction of optimal orders, which may be (but need not be) away from average demand. Minimizing ex post inventory error, in contrast, implies that orders are positively correlated with past demand. Although not mutually exclusive, the two explanations yield different implications about the way that orders adjust over time, which leads us to our third hypothesis.

**HYPOTHESIS 3 (CAUSES FOR THE PULL-TO-CENTER EFFECT).**

**A. ANCHORING AND INSUFFICIENT ADJUSTMENT:** *The retailers' orders for all contracts will adjust, over time, toward the optimal order.*

**B. MINIMIZING EX POST INVENTORY ERROR:** *The retailers' orders will be positively correlated with past demand.*

Our fourth hypothesis deals with the mathematical equivalence of the buyback and revenue-sharing contracts and a possible reason that this equivalence may fail. If these contracts are equivalent, we should not observe any differences in the performance of the two mechanisms in the supplier game, either in the retailers' orders or in the contract parameters in terms of the per-unit cost of units sold and unsold. (In the retailer game, we set the parameters such that the cost of sold units is 9 and the cost of unsold units is 1.) This theoretical equivalence leads to the first part of our fourth hypothesis:

**HYPOTHESIS 4A (MATHEMATICAL EQUIVALENCE OF THE REVENUE-SHARING AND BUYBACK CONTRACTS).** *The retailers' average orders and the per-unit cost of sold and unsold units for the buyback and revenue-sharing contracts will be equal.*

Because the two contracts are mathematically equivalent as long as Equation (3) holds, any differences we observe between them in the laboratory

<sup>6</sup> The previous literature on ordering behavior in the newsvendor problem that we discussed in §2 deals with the wholesale price contract only.

must be due to framing (see Soman 2004 for a review of the literature on framing). In other words, the mathematical equivalence of the contracts may not be immediately apparent to participants. We propose that loss aversion is a potential cause for these differences. In a bilateral monopoly setting, Ho and Zhang (2008) cite loss aversion as the explanation for the differences they observe between the quantity discount and the two-part tariff contract, which are also mathematically equivalent. In our setting, loss aversion is a plausible driver because the wholesale price, which is low under the revenue-sharing contract and high under the buyback contract, is a payment that depends directly on the order amount, and is therefore clearly viewed as a loss for the retailer and a gain for the supplier. But the effect of the buyback rebate and the revenue share depends on the realization of the uncertain demand. Loss aversion is affected by framing because people dislike losses more than they dislike forgone gains (Kahneman and Tversky 1979, Kahneman et al. 1990), and thus the way that we frame demand distribution affects how participants perceive gains and losses.

In the DLOW condition, the retailer is not guaranteed any revenue (because the demand can be as low as 0), so the potential losses to the retailer from paying the wholesale price loom large, whereas potential gains from a rebate ( $b$ ) for unsold units or the forgone losses from having to pay a revenue share ( $r$ ) for the sold units seem less salient. Consequently, the low upfront wholesale price of the revenue-sharing contract may be more effective than the higher rebate of the buyback contract in inducing high retailer order quantities in the retailer game. In the supplier game, however, the situation is exactly reversed, because for the revenue-sharing contract to coordinate, the wholesale price must be below the supplier's cost. Consequently, in the supplier game, the high wholesale price of the buyback contract may be more effective in inducing higher retailer orders than the high revenue share of the revenue-sharing contract.

**HYPOTHESIS 4B.1 (LOSS AVERSION WITH DLOW DEMAND).** *In the DLOW demand condition, average retailer orders will be higher under the revenue-sharing contract in the retailer game and higher under the buyback contract in the supplier game.*

In the DHIGH demand condition, the retailer is guaranteed to sell at least 50 units, and thus the revenue share ( $r$ ) becomes a salient loss for the retailer in the retailer game, but a salient gain for the supplier in the supplier game—at least for those 50 units. This makes the revenue-sharing contract look like it penalizes high orders in the retailer game and rewards high orders in the supplier game. The buyback contract, in contrast, appears to reward high orders in

the retailer game and to penalize them in the supplier game, because a substantial rebate is paid for unsold units.

**HYPOTHESIS 4B.2 (LOSS AVERSION WITH DHIGH DEMAND).** *In the DHIGH demand condition, average retailer orders should be higher under the buyback contract in the retailer game and higher under the revenue-sharing contract in the supplier game.*

When we describe the DHIGH demand as 50 guaranteed units plus a random number of additional units and frame the decision in terms of the number of additional units to order, we move the task into the DLOW frame, and just as in the DLOW condition, the revenue-sharing contract should look more appealing to loss-averse retailers in the retailer game (and less appealing to loss-averse suppliers in the supplier game) than the buyback contract because either paying or receiving the revenue share for the first 50 guaranteed units no longer seems to be part of the decision.

**HYPOTHESIS 4B.3 (LOSS AVERSION WITH DHIGH/LOW DEMAND).** *In the DHIGH/LOW demand condition, average retailer orders will be higher under the revenue-sharing contract in the retailer game and higher under the buyback contract in the supplier game.*

Note that the comparison between the two contracts is the same in the DLOW condition (Hypothesis 4B.1) as in the DHIGH/LOW condition (Hypothesis 4B.3) and is the opposite of that in the DHIGH condition (Hypothesis 4B.2). Whether we observe this reversal is the main test of the loss-aversion and framing explanation.

## 5. Results

### 5.1. Hypothesis 1 Results: Theoretical Benchmarks

In Table 2, we provide descriptive statistics for the retailer order quantities in all treatments, including the average retailer order, standard deviations, and median orders. We used the one-sample Wilcoxon test (Siegel 1956, pp. 75–83) to make the comparisons and conducted the test separately for inexperienced and experienced sessions. The unit of analysis is the average order of an individual subject. Because the decision in the supplier game cannot be fully described by the retailer order induced by the contract, we also provide, in Table 3, descriptive statistics for the contract parameters in the supplier game treatments.<sup>7</sup>

<sup>7</sup> Any order quantity  $q$  can be induced in many different ways, depending on how a particular contract distributes supply chain profits between the two parties. So the decision in the supplier game can be fully described by the retailer order induced ( $q$ ) and the retailer's profit share ( $\lambda$ ), or equivalently and more directly by the cost of the sold and unsold units.

**Table 2 Average Retailer Orders for All Treatments**

Contract	Best-response quantity ( $q^*$ )	Retailer game		Supplier game	
		Inexperienced	Experienced	Inexperienced	Experienced
DLOW: $D \sim U(0, 100)$					
Wholesale	37.5	42.04 (4.91) [41.24]**	41.72 (6.92) [41.64]*	37.44 (3.32) [37.68]	39.59 (5.08) [37.84]
Buyback	75	55.55 (5.89) [55.92]**	57.69 (12.32) [58.86]**	57.05 (10.41) [58.76]**	57.62 (15.48) [53.37]*
Revenue-sharing	75	64.07 (9.20) [61.93]*	69.13 (12.66) [74.13]	49.68 (11.81) [49.93]**	50.60 (17.15) [46.49]*
DHIGH: $D \sim U(50, 150)$					
Wholesale	62.5	81.23 (10.19) [82.77]**	80.49 (10.17) [81.47]**	66.35 (3.57) [64.69]**	66.89 (8.04) [64.60]**
Buyback	125	115.81 (11.56) [113.60]*	106.10 (17.29) [104.45]*	87.00 (16.28) [83.02]**	89.54 (22.69) [85.39]**
Revenue-sharing	125	102.85 (16.17) [104.93]*	104.56 (19.65) [104.23]*	98.46 (11.38) [98.52]**	94.47 (10.18) [96.34]**
DHIGH/LOW: $D = 50 + X, X \sim U(0, 100)$					
Buyback	125	102.71 (9.83) [104.26]**	109.42 (13.72) [108.85]**	100.55 (21.21) [102.30]	97.58 (21.95) [95.77]
Revenue-sharing	125	115.03 (15.20) [109.57]	108.17 (13.36) [110.97]**	102.21 (12.80) [100.84]**	101.20 (19.32) [108.07]**

Notes. For Hypothesis 1,  $H_0: q = q^*$ . Standard deviations are reported in parentheses and median orders in square brackets.

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

The only retailer game treatment for which we cannot reject Hypothesis 1 is the DLOW condition of the revenue-sharing contract in the experienced session. Median orders in the other treatments of the retailer game are generally higher than the optimal orders under wholesale price contracts and are generally lower than the optimal orders under the two coordinating contracts.

The most convenient way to evaluate the buyback and revenue-sharing contracts in the supplier game is by comparing the cost of sold and unsold units. In the buyback contract, the retailer pays  $w_{BB}$  for each unit sold and  $w_{BB} - b$  for each unit unsold. In the revenue-sharing contract, the retailer pays  $w_{RS} + r$  for each unit sold and  $w_{RS}$  for each unit unsold. The two contracts are equivalent if the cost of the sold and unsold units are equal. (In the wholesale price contract, the retailer pays  $w$  for each unit whether or not it is sold.) Therefore, in the supplier game, we can analyze Hypothesis 1 in two ways, through retailer orders induced by the contracts (Table 2), and directly through the contract parameters (Table 3).

Under the wholesale price contract in the DLOW condition, median retailer orders do not differ from

**Table 3 Average Cost of Sold and Unsold Units in Supplier Game Treatments**

Contract	Cost of units sold		Cost of units unsold		
	Inexperienced	Experienced	Optimal (given sold)	Inexperienced	Experienced
DLOW: $D \sim U(0, 100)$					
Wholesale	7.31 (0.81) [7.50]	7.44 (0.88) [7.50]	7.50	7.31 (0.81) [7.50]	7.44 (0.88) [7.50]
Buyback	8.80 (1.62) [9.00]	8.77 (1.19) [9.00]	1.00	2.55 (1.65) [2.00]	2.87 (1.44) [3.00]
Revenue-sharing	8.79 (1.64) [9.00]	8.80 (1.11) [8.72]	1.05	4.41 (2.33) [5.00]	3.48 (1.96) [3.00]
DHIGH: $D = U(50, 150)$					
Wholesale	9.85 (1.32) [10.50]	10.16 (0.94) [10.50]	10.50	9.85 (1.32) [10.50]	10.16 (0.94) [10.50]
Buyback	10.20 (1.07) [10.00]	10.25 (0.91) [10.50]	0.58	3.90 (2.24) [3.50]	4.04 (3.30) [3.00]
Revenue-sharing	9.50 (1.26) [9.90]	10.26 (0.81) [10.00]	0.68	2.89 (1.64) [2.54]	2.39 (1.62) [2.00]
DHIGH/LOW: $D \sim 50 + X, X \sim U(0, 100)$					
Buyback	9.62 (1.67) [10.00]	9.88 (1.41) [10.00]	0.67	3.32 (2.45) [3.00]	3.44 (2.58) [3.00]
Revenue-sharing	9.75 (1.33) [10.00]	10.01 (1.03) [10.00]	0.67	2.38 (1.55) [2.04]	2.59 (2.34) [2.00]

Note. Standard deviations are reported in parentheses and median levels in square brackets.

optimal orders (37.5), and median wholesale prices also are not different from the optimal prices (7.5), consistent with Hypothesis 1. In the DHIGH condition, however, even though median wholesale prices do not differ from the optimal level (10.5), median retailer orders are slightly above optimal, so here Hypothesis 1 is only partially supported.

We can see from Table 2 that retailer orders under both coordinating contracts in the supplier game are significantly below optimal, and Table 3 reveals why. If we look at the cost of unsold units, the column labeled “Optimal (given sold)” tells us what the cost of unsold units should be, given the median cost of sold units. (Because median costs of sold units are very similar in the inexperienced and experienced sessions, we use the average of the two medians for the optimal calculation.) Note that the actual cost of unsold units that our suppliers charge is always substantially higher than it needs to be to coordinate the channel. This is evidence that our suppliers are unwilling to assume enough risk to coordinate the channel.

**5.2. Hypothesis 2 Results: Coordination**

In Table 4 we summarize the results of our tests of Hypothesis 2. The table shows the differences in

INFORMS holds copyright to this article and distributed this copy as a courtesy to the author(s). Additional information, including rights and permission policies, is available at http://journals.informs.org/.

**Table 4** Differences in Median Orders Under Wholesale Price and Coordinating Contracts

Contract	Retailer game		Supplier game	
	Inexperienced	Experienced	Inexperienced	Experienced
DLOW: $D = U(0, 100)$				
Buyback	14.68**	17.22**	21.08**	15.53**
Revenue-sharing	20.69**	32.49**	12.25*	8.65*
DHIGH: $D = U(50, 150)$				
Buyback	30.83**	22.98**	18.33**	20.76**
Revenue-sharing	22.16**	22.76**	33.83**	31.74**
DHIGH/LOW: $D = 50 + X, X \sim U(0, 100)$				
Buyback	41.76**	46.92**	39.80**	33.27**
Revenue-sharing	47.07**	52.53**	38.34**	45.57**

Note. For Hypothesis 2,  $H_0: q_W = q_{BB}; q_W = q_{RS}$ .  
 \* $p < 0.05$ ; \*\* $p < 0.01$ .

median order under the wholesale price contract versus under the coordinating contracts. We use the Mann–Whitney  $U$  test (Siegel 1956, pp. 116–127) to make the comparisons, and we conduct the test separately for inexperienced and experienced sessions. The unit of analysis for this comparison again is the average order of an individual subject. We find support for Hypothesis 2 in all treatments. Both the buyback and revenue-sharing contracts induce higher retailer orders than the wholesale price contract.

### 5.3. Hypotheses 3A and 3B Results: Causes for the “Pull-to-Center” Effect

Because Hypotheses 3A and 3B deal with behavior over time, we test it using a regression model that we fit for each demand condition and contract type. We fit the inexperienced and experienced sessions separately.

$$Q_{i,t} = \text{Intercept} + \beta_t \times (t - 2) + \beta_{D-1} \times \Delta D_{i,t-1} + \eta_i + \varepsilon_{i,t}. \quad (7)$$

In this model, the dependent variable  $Q_{i,t}$  is participant  $i$ 's order in period  $t$ . (In the supplier game it is the automated retailer's order induced by participant  $i$ 's choice of contract parameters in period  $t$ .) The variable  $\Delta D_{i,t-1}$  is the difference between the demand that participant  $i$  observed in period  $t - 1$  and the average demand under the given demand condition (i.e., 50 in the DLOW and 100 in the DHIGH and DHIGH/LOW conditions). The variable  $t - 2$  captures the time trend. Note that there are two error components in the model: one that is independent across all observations,  $\varepsilon_{i,t}$ , and one that is participant specific,  $\eta_i$ . Each error term has a mean of zero and some positive standard deviation. This treatment of the individual effect is known as the random-effects model, and it is used to control for individual heterogeneity. We present regression estimates for Equation (7) in Table A.1 in the appendix. In Table 5, to

make the exposition easier, we present the analysis related to Hypotheses 3A and 3B that includes the estimates of the *Intercept* term and the signs of the two  $\beta$  coefficients.

Because the variable  $\Delta D_{i,t-1}$  captures the difference between the demand observed in the previous period and the average demand, the intercept can be interpreted as the average order at the beginning of the session (at  $t = 2$ ) measured at average values of the demand in period 1. At that point,  $t - 2 = 0$ .

Hypothesis 3A implies that, over time, orders should move toward optimal levels, meaning that if the intercept is above  $q^*$ , the coefficients on  $t - 2$  should be negative; if the intercept is below  $q^*$ , the coefficients on  $t - 2$  should be positive; and if the intercept does not differ from  $q^*$ , the coefficients on  $t - 2$  should be 0. Under the wholesale price contract, the intercept is always above  $q^*$ , and the coefficient on  $t - 2$  is always negative, consistent with Hypothesis 3A. Under the buyback and revenue-sharing contracts, when the intercept is below  $q^*$ , some of the coefficients on  $t - 2$  are negative or not significant, contrary to Hypothesis 3A. In the two cases in which the intercept is not significantly below  $q^*$ , the coefficients on  $t - 2$  are negative, again contrary to Hypothesis 3A. So we find support for Hypothesis 3A under the wholesale price contract, but not under the two coordinating contracts.

Hypothesis 3B implies that the coefficients on  $\Delta D_{i,t-1}$  should be positive. We find that they are in fact positive and significant in each of the retailer game treatments and not significant in any of the supplier game treatments. So we find support for Hypothesis 3B in the retailer game, but not in the supplier game.

The bottom line is that whereas anchoring and adjustment appears to be a reasonably accurate description of behavior under the wholesale price contract, we find no consistent evidence of this behavior under coordinating contracts. In some treatments, average orders move toward the optimum, and in other treatments they move away from it. But in all treatments of the retailer game there is a positive correlation between orders and last-period demand, indicating that the desire to minimize ex post inventory error (the regret from having a mismatch between the order and the demand, as Kremer et al. (2007) point out) plays a role.

### 5.4. Hypotheses 4A–4B.3 Results: Equivalence

Hypotheses 4A–4B.3 deals with the equivalence of the buyback and revenue-sharing contracts. Note that in the retailer game the two contracts are literally equivalent because we set contract parameters to make them so, per Equation (3). In the supplier game, the two contracts can be equivalent if suppliers set contract parameters appropriately, again according to

**Table 5 Results of Analysis of Hypotheses 3A and 3B**

Variable	Retailer game						Supplier game					
	DLOW: $D = U(0, 100)$		DHIGH: $D = U(50, 150)$		DHIGH/LOW: $D = 50 + X,$ $X \sim U(50, 150)$		DLOW: $D = U(0, 100)$		DHIGH: $D = U(50, 150)$		DHIGH/LOW: $D = 50 + X,$ $X \sim U(50, 150)$	
	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced
Wholesale price contract												
Intercept	43.42*	43.01*	84.59*	83.61*			38.33	40.94*	71.02*	71.01*		
	(1.23)	(1.64)	(2.35)	(2.32)			(0.79)	(1.14)	(0.81)	(1.80)		
$t - 2$	–	–	–	–			–	–	–	–		
$\Delta D_{t-1}$	+	+	+	+			0	0	0	0		
Buyback contract												
Intercept	58.02*	58.66*	120.18	102.55*	93.82*	103.74*	60.95*	57.72*	87.87*	85.00*	95.96*	92.10*
	(2.13)	(3.96)	(3.90)	(5.54)	(3.29)	(4.42)	(3.48)	(4.82)	(5.22)	(6.86)	(6.79)	(7.06)
$t - 2$	–	0	–	+	+	+	–	0	0	+	+	+
$\Delta D_{t-1}$	+	+	+	+	+	+	0	0	0	0	0	0
Revenue-sharing contract												
Intercept	65.87*	66.99*	100.13*	102.17*	106.19*	119.15	48.44*	48.30*	94.54*	88.67*	101.28*	99.00*
	(3.03)	(4.05)	(5.22)	(6.26)	(4.32)	(4.89)	(3.82)	(5.46)	(3.18)	(4.12)	(6.16)	(4.21)
$t - 2$	–	+	+	+	+	–	+	+	+	+	0	+
$\Delta D_{t-1}$	+	+	+	+	+	+	0	0	0	0	0	0

Note. Intercept standard deviations are reported in parentheses.  
 \*Intercept  $\neq 0$ ,  $p < 0.05$ ; –,  $\beta < 0$ ,  $p < 0.05$ ; +,  $\beta > 0$ ,  $p < 0.05$ ; 0,  $p > 0.05$ .

Equation (3). We compare the two coordinating contracts in three ways. First, in both the retailer and the supplier games we compare the average retailer orders ( $Q$ ). Second, because the average retailer order does not fully describe the supplier game contract, we also compare the average cost of sold and unsold units in that game. Together,  $Q$  and the cost of sold and unsold units fully describe the supplier game contracts. To test Hypotheses 4A–4B.3, we then fit three models for the buyback and revenue-sharing treatments in the retailer and supplier games. The first model looks at the retailer’s order in both games:

$$Q_{i,t} = \text{Intercept} + \beta_{RS} \times RS + \beta_t \times (t - 2) + \beta_{RS \times t} \times [RS \times (t - 2)] + \beta_{D-1} \times \Delta D_{i,t-1} + \eta_i + \varepsilon_{i,t}. \quad (8)$$

The next two models look at the cost of sold and unsold units in the supplier game only:

$$\begin{aligned} \text{Sold}_{i,t} &= \text{Intercept} + \beta_{RS} \times RS + \beta_t \times (t - 2) + \beta_{RS \times t} \times [RS \times (t - 2)] + \beta_{D-1} \times \Delta D_{i,t-1} + \eta_i + \varepsilon_{i,t}, \\ \text{Unsold}_{i,t} &= \text{Intercept} + \beta_{RS} \times RS + \beta_t \times (t - 2) + \beta_{RS \times t} \times [RS \times (t - 2)] + \beta_{D-1} \times \Delta D_{i,t-1} + \eta_i + \varepsilon_{i,t}, \end{aligned} \quad (9)$$

where

$$\begin{aligned} \text{Sold}_{i,t} &= w_{BB} && \text{for buyback} \\ &= w_{RS} + r && \text{for revenue-sharing,} \\ \text{Unsold}_{i,t} &= w_{BB} - b && \text{for buyback} \\ &= w_{RS} && \text{for revenue-sharing.} \end{aligned}$$

We estimate these models for each demand condition and for inexperienced and experienced sessions separately. Note that the independent variables we use in Equations (8) and (9) are the same as in Equation (7), but we include an additional variable,  $RS$ , to measure the differences between the buyback and revenue-sharing contracts. This variable is the focus of our analysis. The variable  $RS$  takes a value of 1 for revenue-sharing treatments and 0 for the buyback treatment, so  $\beta_{RS}$  measures the differences between the two contracts (in terms of the dependent variables). Because we know from the estimates of Equation (7) that time trends under the buyback and revenue-sharing contracts are not always the same, we added the interaction variable  $[RS \times (t - 2)]$  to control for the differences in the way that participants adjust their decisions over time under the two contracts. Because the coefficients on  $\Delta D_{i,t-1}$  are very similar for the buyback and the revenue-sharing contracts in estimating Equation (7), we do not need a similar interaction term between  $RS$  and  $\Delta D_{i,t-1}$ .<sup>8</sup> We report  $\beta_{RS}$  estimates in Table 6 and the full results of estimates of Equations (8) and (9) in Tables A.2–A.5 in the appendix.

Hypothesis 4A predicts that none of the  $RS$  coefficients should be significant. We do, however, observe some significant  $RS$  coefficients when the dependent variable is  $Q_{i,t}$ . Also, one of the  $RS$  coefficients is significant when the dependent variable is  $\text{Sold}_{i,t}$ , and

<sup>8</sup> However, when we add such an interaction term to the model, the interaction variables are not significantly different from 0, and the rest of the estimates remain virtually unchanged.

INFORMS holds copyright to this article and distributed this copy as a courtesy to the author(s). Additional information, including rights and permission policies, is available at http://journals.informs.org/.

**Table 6** Estimates of  $\beta_{RS}$

Contract	Retailer game				Supplier game			
	$Q_{i,t}$		$Q_{i,t}$		Sold		Unsold	
	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced
DLOW: $D \sim U(0, 100)$	7.40* (3.69)	8.18 (5.67)	-9.63* (4.75)	-9.19 (7.29)	0.02 (0.59)	0.14 (0.45)	2.34** (0.78)	1.13 (0.70)
DHIGH: $D \sim U(50, 150)$	-19.89** (6.52)	-0.79 (8.36)	14.11* (5.68)	3.88 (8.21)	-0.57* (0.28)	0.03 (0.32)	-1.19 (0.72)	-1.66 (1.04)
DHIGH/LOW: $D = 50 + X, X \sim U(0, 100)$	25.32** (5.89)	2.43 (6.17)	9.15 (9.35)	3.01 (7.98)	0.16 (0.58)	0.38 (0.44)	-1.29 (0.74)	-1.80 (0.97)

Notes. For Hypotheses 4A–4B.3,  $H_0: \beta_{RS} = 0$ . Standard deviations are reported in parentheses.

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

another is significant when the dependent variable is  $Unsold_{i,t}$ . We conclude that the data are generally not consistent with Hypothesis 4A, because we observe some differences between the two coordinating contracts.

The next question is whether the differences we observe between the coordinating contracts are consistent with loss aversion, and whether the differences are affected by the framing. Hypotheses 4B.1 and 4B.3 predict that in the retailer game when the dependent variable is  $Q_{i,t}$ , the RS coefficient will be positive in the DLOW and DHIGH/LOW conditions; Hypothesis 4B.2 predicts that it will be negative in the DHIGH condition. The sign of the RS coefficient in the supplier game thus should be the opposite of its sign in the retailer game. Because in every case in which the RS coefficient is significantly different from 0 its sign is consistent with the predictions of Hypotheses 4B.1–4B.3, we find that the data offer some support for these hypotheses. The strongest evidence we have to offer in support of the loss-aversion hypothesis is that in the retailer game, the sign of the RS coefficient switches from being strongly negative in the inexperienced session of the DHIGH condition to strongly positive in the inexperienced session of the DHIGH/LOW condition. The only difference between the inexperienced sessions of those two treatments is how we framed the demand distribution, and we therefore conclude that framing (in particular the effect of framing on loss aversion) is the only possible explanation.

Another regularity we observe is that none of the RS coefficients is significant in the experienced sessions. This implies that to the extent that loss aversion affects behavior, it tends to lessen, and usually disappear, with experience in our setting. In other words, as participants gain experience, they are less affected by framing. So we find some evidence consistent with loss aversion (Hypotheses 4B.1–4B.3) but also find that the effect of loss aversion appears to wear off over time, to the point that when participants play for

the second time, there is generally no detectable difference between the two coordinating contracts. This is consistent with Hypothesis 4A.

## 6. Summary, Limitations, and Managerial Implications

### 6.1. Summary of Results

In this laboratory study we compare the performance of the wholesale price contract and two types of coordinating risk-sharing contracts, the buyback and revenue-sharing contracts. We first look at how retailers respond to different mechanisms; we then examine suppliers' willingness and ability to take advantage of coordinating contracts.

We find that, consistent with earlier studies, retailers on average place orders that are between the profit-maximizing order and the average demand. In the context of wholesale price contracts, average retailer orders are higher than the expected-profit maximizing benchmark, and they adjust in the direction of the optimal order over time. This initial overordering behavior causes wholesale price contracts to perform better in our laboratory setting than in theory.<sup>9</sup> Coordinating contracts induce higher retailer orders than do wholesale price contracts, but they fall short of the channel-optimal solution, so they perform worse in the laboratory than in theory.

We also find that suppliers, when interacting with computerized retailers that are programmed to respond optimally, quickly find the wholesale price that maximizes their expected profit. We do not, however, observe this under either of the coordinating contracts, where suppliers induce retailer orders that

<sup>9</sup> The fact that wholesale price contracts perform better than they should in theory is because the optimal order in our experiments is below average demand. When the optimal order is above average demand, actual orders are below the optimal level (see Schweitzer and Cachon 2000, Bolton and Katok 2008). So we cannot conclude that wholesale contracts generally perform better in the laboratory than in theory.

are significantly below the channel-optimal level. We offer two explanations, the first related to bounded rationality, and the second related to preferences. One explanation is that suppliers do not face any demand risk under the wholesale price contract, but they do under coordinating contracts, so they perform better under the wholesale price contracts because deterministic problems are generally easier to solve than are stochastic problems. A related issue is that suppliers have to select two different contract parameters under coordinating contracts, which is a more complex task than selecting a single parameter under the wholesale price contract. A review of the contract parameters that our suppliers chose points to a second explanation. Suppliers systematically set the cost of unsold units significantly above the channel-optimal level, given the cost they set for sold units. Thus, suppliers do not take enough risk to coordinate the channel. With human retailers, who are likely to place orders that are even lower than our computerized retailers, the effectiveness of coordinating contracts may well be even worse.

Our study is the first to examine coordinating contracts in the laboratory in an environment with stochastic customer demand. We find that particularly for coordinating contracts, the changes in retailer behavior over time are more consistent with the preference for minimizing *ex post* inventory error than with the anchoring and adjustment heuristic. This is evidenced by the fact that in every treatment of the retailer game, retailers placed orders that were positively correlated with last-period demand. Suppliers, though, did not choose contract parameters in a way that caused the retailer orders they induced to be positively correlated with past demand.

We find that the two mathematically equivalent risk-sharing contracts do not initially induce identical retailer or supplier behavior in the laboratory; however, the observed differences tend to decrease and disappear with experience. We offer a framing explanation for this initial lack of equivalence, specifically, that participants' perception of the demand distribution is an influential factor. The buyback contract emphasizes the benefit of placing higher orders, so it is more effective when the demand distribution is framed in terms of a minimum amount plus the possibility of an upside. The revenue-sharing contract, which emphasizes low upfront cost, is more effective when the demand distribution frames the decision as one with less of an upside and a serious potential downside. The results from our study of the DHIGH demand distribution with the DLOW decision frame demonstrate that although the framing concept is useful for explaining initial differences in behavior, differences due to framing tend to disappear with experience.

## 6.2. Limitations and Directions for Future Research

Our study is subject to two main limitations that point toward fruitful directions for future research. The first has to do with the subject pool. We used a subject pool that is common in experimental economics (see Holt 1995, Kagel and Roth 1995) comprising students, mostly undergraduates, recruited through advertisements offering an opportunity to earn cash. Our subject pool is representative of the larger student population at Penn State in terms of gender, majors, and, to the extent we are able to determine, ethnicity. A number of studies have found no difference between the performance of students and professionals in laboratory experiments (see, e.g., Plott 1987, Ball and Cech 1996, Katok et al. 2008). However, in the context of the newsvendor problem, Bolton et al. (2008) investigated the effect of learning and found that instruction regarding how to solve the newsvendor problem was highly effective with students but had almost no effect on managers. Clearly, the effect of the subject pool on the outcome of operations management experiments is a complicated and important question, deserving of future study.

The second limitation is that we study the behavior of retailers and suppliers separately—our retailers and suppliers do not interact. This design was intentional, because we wanted to understand individual decision making unaffected by social utility considerations such as preferences for fairness (Fehr and Schmidt 1999, Bolton and Ockenfels 2000, Cui et al. 2007). In that respect, our study represents an intermediate step, because it offers a standard for comparison that can be used to separate the effect of social preferences from other individual decision-making biases. Whereas Keser and Paleologo (2004) reported that in their study, which included two human players, wholesale prices were consistently below optimal, we found that in our study suppliers quickly identified profit-maximizing wholesale price contracts. Based on our findings, we can say with a fair degree of confidence that Keser and Paleologo's (2004) result is likely a consequence of social preferences, which was not their original conclusion.

Real contractual arrangements, however, are negotiated by human participants on both sides, so a better understanding of how these contracts compare in reality would be gained through a study that includes the interaction of human retailers and suppliers. Wu (2009) follows this direction, extending our experimental design to investigate the strategic interactions between supply chain partners under buyback and revenue-sharing contracts. In Wu's (2009) study, retailers are constrained to place the optimal order or to punish the supplier by either ordering a quantity of 0 or ordering the minimum possible demand amount.

Wu's (2009) main finding is that in that environment participants learn, over time, to negotiate contracts that are more efficient than the contracts in our supplier game, while dividing expected profits in a more equitable manner.

### 6.3. Managerial Implications

Our results indicate that decision-support tools for contract design could increase the effectiveness of contractual arrangements for both suppliers and retailers. From the retailers' perspective, tools are needed to counteract the demand-following behavior that results from participants' trying to minimize ex post inventory error. It is this behavior that decreases the effectiveness of coordinating contracts by making retailers less responsive to economic incentives (see Kremer et al. 2007). From the suppliers' perspective, decision-support tools that help set contract parameters properly may well go a long way to increase not only the total supply chain efficiency, but also the suppliers' profit. When buyback contracts are prevalent in an industry, the natural tendency of contract designers may be to set both the wholesale price and the rebate too low. Alternatively, a lack of understanding of contracting mechanisms may lead to such obviously suboptimal contracts as the offering of full rebates,

as is the standard in the pharmaceutical industry, for example. Although such contracts may be rational in the face of retailer underordering, they are sure to lead to over-ordering. Thus, effective decision-support tools for both retailers and suppliers offer promise for decreasing waste and increasing profitability.

## 7. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at <http://mansci.journal.informs.org/>.

### Acknowledgments

This research was funded in part by a National Science Foundation Doctoral Dissertation Grant (SES-0412900) and by the Smeal College of Business Doctoral Dissertation Award, Pennsylvania State University. The authors thank Axel Ockenfels and the Deutsche Forschungsgemeinschaft for financial support through the Leibniz Program. Katok gratefully acknowledges support from the Smeal College of Business and the Center for Supply Chain Research at Pennsylvania State's Smeal College of Business through the Smeal Summer Grants Program. The authors thank Kay-Yut Chen, Karen Donohue, Mirko Kremer, and Enno Siemsen for helpful suggestions and feedback, and Lynn Hand for copyediting. All remaining errors are their own.

## Appendix

**Table A.1** Tests of Hypotheses 3A and 3B: Estimates of Equation (7)

Variable	Retailer game						Supplier game					
	DLOW: $D = U(0, 100)$		DHIGH: $D = U(50, 150)$		DHIGH/LOW: $D = 50 + X,$ $X \sim U(50, 150)$		DLOW: $D = U(0, 100)$		DHIGH: $D = U(50, 150)$		DHIGH/LOW: $D = 50 + X,$ $X \sim U(50, 150)$	
	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced
Wholesale price contract												
Intercept	43.42** (1.23)	43.01** (1.64)	84.59* (2.35)	83.61* (2.32)			38.33 (0.79)	40.94* (1.14)	71.02* (0.81)	71.01* (1.80)		
$t - 2$	-0.028** (0.0109)	-0.025** (0.0102)	-0.069** (0.0109)	-0.068** (0.0090)			-0.013** (0.0045)	-0.023** (0.0041)	-0.097** (0.0053)	-0.072** (0.0054)		
$\Delta D_{t-1}$	0.0948** (0.0108)	0.0919** (0.0101)	0.1079** (0.0108)	0.0932** (0.0090)			0.0006 (0.0044)	0.0024 (0.0040)	-0.0024 (0.0053)	-0.0047 (0.0054)		
Buyback contract												
Intercept	58.02* (2.13)	58.66* (3.96)	120.18 (3.90)	102.55* (5.54)	93.82* (3.29)	103.74* (4.42)	60.95* (3.48)	57.72* (4.82)	87.87* (5.22)	85.00* (6.86)	95.96* (6.79)	92.10* (7.06)
$t - 2$	-0.0510** (0.0211)	-0.0190 (0.0154)	-0.069** (0.0193)	0.0751** (0.0170)	0.1829** (0.0205)	0.1144** (0.0144)	-0.0915** (0.0190)	-0.0020 (0.0125)	-0.0180 (0.0139)	0.0972** (0.0114)	0.0881** (0.0157)	0.1054** (0.0194)
$\Delta D_{t-1}$	0.1163** (0.0217)	0.1075** (0.0158)	0.0648** (0.0198)	0.0717** (0.0174)	0.0529* (0.0213)	0.0263* (0.0149)	0.0065 (0.0195)	-0.0153 (0.0128)	-0.0024 (0.0143)	-0.0017 (0.0117)	-0.0073 (0.0163)	-0.0008 (0.0201)
Revenue-sharing contract												
Intercept	65.87* (3.03)	66.99* (4.05)	100.13* (5.22)	102.17* (6.26)	106.19* (4.32)	119.15 (4.89)	48.44* (3.82)	48.30* (5.46)	94.54* (3.18)	88.67* (4.12)	101.28* (6.16)	99.00* (4.21)
$t - 2$	-0.0364** (0.0169)	0.0441** (0.0121)	0.0578** (0.0177)	0.0506** (0.0131)	0.0416* (0.01837)	-0.0823** (0.0174)	0.0280* (0.0142)	0.0487** (0.0112)	0.0590** (0.0151)	0.1734** (0.0186)	-0.0081 (0.0153)	0.0624** (0.0187)
$\Delta D_{t-1}$	0.0651** (0.0174)	0.0912** (0.0125)	0.0828** (0.0181)	0.0245** (0.0135)	0.0351* (0.0188)	0.057** (0.0179)	0.0144 (0.0146)	0.0102 (0.0115)	0.0051 (0.0155)	0.0221 (0.0191)	0.0191 (0.0157)	0.0265 (0.0193)

Note. Coefficient estimates and standard errors are reported in parentheses.

\* $p < 0.10$ ; \*\* $p < 0.05$ .

**Table A.2 Tests of Hypotheses 4A–4B.3: Estimates of Equation (8) for Retailer Game**

Variable	DLOW: $D \sim U(0, 100)$		DHIGH: $D \sim U(50, 150)$		DHIGH/LOW: $D = 50 + X, X \sim U(0, 100)$	
	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced
<i>Intercept</i>	58.25** (2.61)	58.73** (4.01)	120.10** (4.61)	102.75** (5.91)	93.85** (4.16)	103.79** (4.37)
<i>RS</i>	7.40* (3.69)	8.18 (5.67)	-19.89** (6.51)	-0.78 (0.01)	25.32** (5.89)	2.43 (6.18)
$t - 2$	-0.054** (0.0190)	-0.0203 (0.0127)	-0.0676** (0.0183)	0.0710** (0.0151)	0.1824** (0.0189)	0.1133** (0.0164)
$RS \times (t - 2)$	0.0234 (0.0266)	0.0658** (0.0193)	0.1239** (0.0258)	-0.0164 (0.0211)	-0.2651** (0.0269)	-0.0724** (0.2339)
$\Delta D_{t-1}$	0.0908** (0.0139)	0.0994** (0.0101)	0.0738** (0.0134)	0.0481* (0.0110)	0.0548** (0.0139)	0.0308* (0.0120)

Note. Coefficient estimates and standard errors are reported in parentheses.

\* $p < 0.10$ ; \*\* $p < 0.05$ .

**Table A.3 Tests of Hypotheses 4A–4B.3: Estimates of Equation (8) for Supplier Game**

Variable	DLOW: $D \sim U(0, 100)$		DHIGH: $D \sim U(50, 150)$		DHIGH/LOW: $D = 50 + X, X \sim U(0, 100)$	
	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced
<i>Intercept</i>	59.49** (3.54)	57.61** (5.15)	84.14** (4.24)	84.90** (5.67)	96.37** (5.64)	92.44** (6.61)
<i>RS</i>	-9.64* (4.75)	-9.19 (7.29)	14.11* (5.69)	3.88 (8.21)	9.15 (9.35)	3.01 (7.98)
$t - 2$	0.1167** (0.0234)	0.0463** (0.0165)	0.0792** (0.0202)	0.0721** (0.0211)	-0.1108** (0.0247)	-0.0235 (0.0245)
$RS \times (t - 2)$	-0.0898** (0.0166)	0.0002 (0.0118)	-0.0187 (0.0144)	0.0991** (0.0147)	0.1021** (0.0173)	0.0836** (0.0172)
$\Delta D_{t-1}$	0.0103 (0.0122)	-0.0026 (0.0086)	0.0023 (0.0105)	0.0096 (0.0110)	0.0095 (0.0128)	0.0103 (0.0127)

Note. Coefficient estimates and standard errors are reported in parentheses.

\* $p < 0.10$ ; \*\* $p < 0.05$ .

**Table A.4 Tests of Hypotheses 4A–4B.3: Estimates of Equation (9) for Units Sold**

Variable	DLOW: $D \sim U(0, 100)$		DHIGH: $D \sim U(50, 150)$		DHIGH/LOW: $D = 50 + X, X \sim U(0, 100)$	
	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced
<i>Intercept</i>	8.17** (0.42)	8.61** (0.32)	9.66** (0.20)	10.15** (0.22)	9.18** (0.41)	9.37** (0.31)
<i>RS</i>	0.02 (0.59)	0.14 (0.45)	-0.57* (0.28)	0.03 (0.32)	0.16 (0.58)	0.38 (0.44)
$t - 2$	0.0128** (0.0012)	0.0042** (0.0007)	0.0112** (0.0010)	0.0022** (0.0005)	0.0093** (0.0009)	0.0109** (0.0009)
$RS \times (t - 2)$	-0.0009 (0.0017)	-0.0023* (0.0010)	-0.0025 (0.0015)	-0.0003 (0.0007)	-0.0008 (0.0013)	-0.0054** (0.0012)
$\Delta D_{t-1}$	-0.0005 (0.0009)	0.0009 (0.0005)	0.0005 (0.0008)	0.0010* (0.0004)	0.0007 (0.0007)	0.0004 (0.0006)

Notes. The dependent variable is *units sold*. Coefficient estimates and standard errors are reported in parentheses.

\* $p < 0.10$ ; \*\* $p < 0.05$ .

**Table A.5 Tests of Hypotheses 4A–4B.3: Estimates of Equation (9) for Units Unsold**

Variable	DLOW: $D \sim U(0, 100)$		DHIGH: $D \sim U(50, 150)$		DHIGH/LOW: $D = 50 + X, X \sim U(0, 100)$	
	Inexperienced	Experienced	Inexperienced	Experienced	Inexperienced	Experienced
<i>Intercept</i>	2.54** (0.55)	2.78** (0.49)	4.70** (0.51)	4.79** (0.72)	3.97** (0.53)	4.72** (0.69)
<i>RS</i>	2.34** (0.78)	1.13 (0.70)	-1.19 (0.72)	-1.66 (1.04)	-1.29 (0.74)	-1.80 (0.97)
<i>t - 2</i>	-0.0014 (0.0014)	0.0009 (0.0010)	-0.0166** (0.0014)	-0.0155** (0.0014)	-0.0130** (0.0015)	-0.0259** (0.0014)
<i>RS × (t - 2)</i>	-0.0099** (0.0019)	-0.0109** (0.0013)	0.0039* (0.0019)	0.0000 (0.0020)	0.0066** (0.0021)	0.0192** (0.0020)
$\Delta D_{t-1}$	-0.0012 (0.0010)	-0.0008 (0.0007)	0.0002 (0.0010)	-0.0001 (0.0010)	-0.0023* (0.0011)	-0.0002 (0.0010)

Notes. The dependent variable is *units unsold*. Coefficient estimates and standard errors are reported in parentheses.

\* $p < 0.10$ ; \*\* $p < 0.05$ .

## References

- Balachandran, K. R., S. Radhakrishnan. 2005. Quality implications of warranties in a supply chain. *Management Sci.* **51**(8) 1266–1277.
- Ball, S. B., P. Cech. 1996. Subject pool choice and treatment effects in economic laboratory research. *Experiment. Econom.* **6** 239–292.
- Benzion, U., Y. Cohen, R. Peled, T. Shavit. 2008. Decision-making and the newsvendor problem: An experimental study. *J. Oper. Res. Soc.* **59**(9) 1281–1287.
- Bolton, G., E. Katok. 2008. Learning-by-doing in the newsvendor problem: A laboratory investigation. *Manufacturing Service Oper. Management* **10**(3) 519–538.
- Bolton, G., A. Ockenfels. 2000. A theory of equity, reciprocity, and competition. *Amer. Econom. Rev.* **90**(1) 166–193.
- Bolton, G., A. Ockenfels, U. Thonemann. 2008. Managers and students as newsvendors: How out-of-task experience matters. Working paper, University of Cologne, Cologne, Germany.
- Bostian, A. J. A., C. A. Holt, A. M. Smith. 2008. Newsvendor “pull-to-center” effect: Adaptive learning in a laboratory experiment. *Manufacturing Service Oper. Management* **10**(4) 590–608.
- Burnetas, A., P. Ritchken. 2005. Option pricing with downward-sloping demand curves: The case of supply chain options. *Management Sci.* **51**(4) 566–580.
- Cachon, G. P. 2003. Supply chain coordination with contracts. A. G. de Kok, S. C. Graves, eds. *Handbooks in Operations Research and Management Science: Supply Chain Management*, Vol. 11. Elsevier B.V., Amsterdam, 229–340.
- Cachon, G. P., M. A. Lariviere. 2005. Supply chain coordination with revenue sharing contracts: Strengths and limitations. *Management Sci.* **51**(1) 30–44.
- Choi, K., J. G. Dai, J. Song. 2004. On measuring supplier performance under vendor-managed inventory programs in capacitated supply chains. *Manufacturing Service Oper. Management* **6**(1) 53–72.
- Corbett, C. J., J. C. Fransoo. 2007. Entrepreneurs and newsvendors: Do small businesses follow the newsvendor logic when making inventory decisions? Working paper, University of California, Los Angeles.
- Corbett, C. J., D. Zhou, C. S. Tang. 2004. Designing supply contracts: Contract type and information asymmetry. *Management Sci.* **50**(4) 550–559.
- Cui, T. H., J. S. Raju, Z. J. Zhang. 2007. Fairness and channel coordination. *Management Sci.* **53**(8) 1303–1314.
- De Bruyn, A., G. E. Bolton. 2008. Estimating the influence of fairness on bargaining behavior. *Management Sci.* **54**(10) 1774–1791.
- Donohue, K. 2000. Efficient supply contracts for fashion goods with forecast updating and two production modes. *Management Sci.* **46**(11) 1397–1411.
- Fehr, E., K. M. Schmidt. 1999. A theory of fairness, competition and cooperation. *Quart. J. Econom.* **114**(3) 817–868.
- Ferguson, M. E., V. D. Guide, G. C. Souza. 2006. Supply chain coordination for false failure returns. *Manufacturing Service Oper. Management* **8**(4) 376–393.
- Ho, T., J. Zhang. 2008. Designing pricing contracts for boundedly rational customers: Does the framing of the fixed fee matter? *Management Sci.* **54**(4) 686–700.
- Holt, C. A. 1995. Industrial organization: A survey of laboratory results. J. Kagel, A. Roth, eds. *Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ, 349–443.
- Kagel, J. H., A. E. Roth, eds. 1995. *Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ.
- Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47** 263–291.
- Kahneman, D., J. L. Knetsch, R. H. Thaler. 1990. Experimental tests of the endowment effect and the Coase theorem. *J. Political Econom.* **98** 1325–1348.
- Kamrad, B., A. Siddique. 2004. Supply contracts, profit sharing, switching, and reaction options. *Management Sci.* **50**(1) 64–82.
- Katok, E., V. Pavlov. 2009. Fairness and coordination failures in supply chain contracts. Working paper, Pennsylvania State University, University Park.
- Katok, E., D. Thomas, A. Davis. 2008. Inventory service level agreements as coordination mechanisms: The effect of review periods. *Manufacturing Service Oper. Management* **10**(4) 609–624.
- Keser, C., G. Paleologo. 2004. Experimental investigation of retailer-supplier contracts: The wholesale price contract. Working paper, IBM Research, New York.
- Kleindorfer, P. R., D. J. Wu. 2003. Integrating long- and short-term contracting via business-to-business exchanges for capital-intensive industries. *Management Sci.* **49**(11) 1597–1615.
- Kremer, M., S. Minner, L. Van Wassenhove. 2007. Regret and framing effects in the newsvendor model—The impact of framing and complexity. Working paper, INSEAD, France.
- Lee, H., S. Whang. 1999. Decentralized multi-echelon supply chains: Incentives and information. *Management Sci.* **45**(5) 633–640.
- Loch, C. H., Y. Wu. 2008. Social preferences and supply chain performance: An experimental study. *Management Sci.* **54**(11) 1835–1849.

- Lurie, N. H., J. M. Swaminathan. 2009. Is timely information always better? The effect of feedback frequency on decision making. *Organ. Behav. Human Decision Processes* **108**(2) 315–329.
- Padmanabhan, V., I. P. L. Png. 1995. Returns policies: Make money by making good. *Sloan Management Rev.* **37**(1) 65–72.
- Plott, C. R. 1987. Dimensions of parallelism: Some policy applications of experimental methods. A. E. Roth, ed. *Laboratory Experimentation in Economics: Six Points of View*. Cambridge University Press, New York, 193–232.
- Rudi, N., S. Kapur, D. F. Pyke. 2001. A two-location inventory model with transshipment and local decision making. *Management Sci.* **47**(12) 1668–1680.
- Schweitzer, M., G. Cachon. 2000. Decision bias in the newsvendor problem: Experimental evidence. *Management Sci.* **46**(3) 404–420.
- Siegel, S. 1956. *Nonparametric Statistics for Behavioral Sciences*. McGraw-Hill, New York.
- Soman, D. 2004. Framing, loss aversion, and mental accounting. D. J. Hoehler, N. Harvey, eds. *Blackwell Handbook of Judgment and Decision Making*. Blackwell Publishing, Boston, 379–398.
- Spengler, J. J. 1950. Vertical integration and antitrust policy. *J. Political Econom.* **58**(4) 347–352.
- Wu, D. J., P. R. Kleindorfer. 2005. Competitive options, supply contracting, and electronic markets. *Management Sci.* **51**(3) 452–466.
- Wu, Y. D. 2009. The impact of long-term relationship on supply chain contracts: A laboratory study. Working paper, University of Kansas, Lawrence.